AD_____

Award Number: DAMD17-99-1-9294

TITLE: Automated Spot Mammography for Improved Imaging of Dense Breasts

PRINCIPAL INVESTIGATOR: Mitchell M. Goodsitt, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, MI 48109-1274

REPORT DATE: October 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

BEST AVAILABLE COPY

20040413 059

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2003	Annual (15 Sep 2002 - 14 Sep 2003)

4. TITLE AND SUBTITLE
Automated Spot Mammography for Improved Imaging of Dense Breasts

5. FUNDING NUMBERS
DAMD17-99-1-9294

6. AUTHOR(S)
Mitchell M. Goodsitt, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, MI 48109-1274

E-Mail: goodsitt@umich.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates: ALL DTIC reproductions will be in black and white

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

We are developing an automated stereo spot mammography technique to improve imaging of lesions within dense breast tissue. During the fourth year of this project, our work was devoted primarily to: 1) completing our observer study comparing the suspicious spot regions selected by radiologists to those detected by a computer (CAD), 2) performing a study comparing the radiologist selected regions and CAD selected regions to true regions determined by a radiologist from mammograms, biopsy images and pathology results, 3) defining the geometric requirements for the automated collimator to insure the suspicious region in a full-field digital mammogram is adequately covered with stereo spot image acquisition, and 4) obtaining our first images of a modular breast phantom that was manufactured for our spot mammography experiments. The images of the phantom showed it did not satisfy our design requirements with respect to dense regions overlapping the simulated masses, and with respect to the phantom producing an image texture similar to that of an actual mammogram. We therefore were not able to complete our planned phantom experiments. The manufacturer has promised to build us a correct phantom that we will employ in experiments during an extended year of the project.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
44

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

(3) TABLE OF CONTENTS

(4) INTRODUCTION

The purpose of our project is to develop a novel technique for improved imaging of suspicious dense regions within full-field digital mammograms. The basic idea is to automatically detect any large dense region within a whole breast digital mammogram, and to take a second 3D stereo digital mammogram of only that region using automated spot collimation, manual x-ray tube shift for acquiring the left- and right-eye stereo images, and stereoscopic viewing of the images on a stereo display system. This second "stereo spot mammogram" is taken within seconds of the full-field mammogram while the breast is maintained in the same compression. Since only the dense region is exposed, detector saturation will not be an issue. Restriction of the x-ray beam to a small area will limit the amount of breast tissue exposed to radiation and will decrease the volume of tissue that scatters x-rays thereby improving the contrast between a lesion and its surrounding dense breast tissue. Stereoscopic image acquisition and display will enable radiologists to view the suspicious regions in 3-D. This will reduce the tissue superposition problem inherent in conventional single projection mammography, wherein lesions can sometimes be camouflaged by overlying and/or underlying tissues, and normal tissues can sometimes overlap producing an object in the image with a lesion-like appearance. Furthermore, in comparison with conventional spot compression, the automated technique should produce more accurate spot imaging of suspicious regions because it eliminates the need for the repositioning and recompression of the breast for the spot images, and the "spot" location is determined by computer analysis of the digital full-breast image rather than estimated by the technologist by eye from a radiograph. In brief, the project entails development of software to analyze digital mammograms and automatically delineate suspicious dense tissue regions, and the development of hardware to automatically collimate the x-ray beam to those regions and to enable stereo spot imaging of those regions while maintaining the breast in the same position and compression as for the original full-breast digital mammogram.

(5) BODY

We applied for a 1 year no-cost time extension for our project to assure adequate completion of all of the tasks in our Modified Statement of Work (dated January 7,2002). We received notice of the approval of our request for this extension on October 15, 2003. Therefore the following is a fourth year rather than a final report. During the fourth year of this project (September 15, 2002 - September 14, 2003), the following was accomplished relative to the tasks in the approved modified statement of work.

A) Task 1: Develop software to recognize and delineate dense breast regions in full-breast mammograms.

This year, we completed our observer study comparing the regions selected by radiologists for spot imaging with those selected by our in-house developed computer aided detection (CAD) program. We developed a new overlap index to assess the agreement between the regions. This index is the intersection between a spot region of interest (ROI) selected by a radiologist and the spot ROI selected by the CAD program divided by the area of the CAD ROI. This index is thus 100% whenever the CAD region is completely contained within the radiologist selected region, representing complete agreement. We believe this index more accurately depicts the amount of agreement between the ROIs than our previous index, which was the ratio of the intersection area

of the radiologist's and CAD's ROIs divided by the union of the areas. The reason is that even if the CAD ROI were completely contained within the radiologist's ROI, the previous index could be very small if the radiologist-selected region were fairly large.

We also developed software to display true regions of interest (TROIs) for our database of digitized mammograms. The TROIs are regions traced by a radiologist of true masses from analyses of the original film mammograms, biopsy images and pathology results. Of the 200 images that were evaluated by each radiologist in our observer study, 83 contained a single true mass, 13 contained two true masses, and 2 contained three true masses. Thus, there was a total of 115 (= 83 +13 x 2+2 x 3) true masses. We computed a separate overlap index between the TROIs and the radiologist's ROIs and between the TROIs and the CAD's ROIs. This index was defined to be equal to the area of intersection divided by the smaller area. This definition yields a value of 100% if the radiologist's or CAD's ROI is completely contained within the TROI or if the TROI is completely contained within the CAD's or radiologist's ROI. It is an indication of the degree to which there is a "hit" between the radiologist's or CAD's ROI and the TROI. In addition, we computed the number of "hits" by determining the number of cases in which the overlap index with the TROI was greater than or equal to a threshold value of 25%. The 25% value was considered a reasonable threshold for indicating agreement in ROIs. For the computation of the overlap indices between radiologist and CAD regions, all possible pairings of the selected regions were considered, and the indices were ordered from largest to smallest. For the radiologist vs. TROI and CAD vs. TROI cases, we computed the largest overlap indices individually for the first TROI, second TROI, and third TROI in the image

Five MQSA qualified radiologists performed the observer study using our TRACEIMAGE software to trace up to 3 regions for spot imaging in each digitized mammogram. The images could be displayed in both 200-micron and 400- micron resolution as described in last year's report. Each radiologist performed the study twice with the reading session separated by a minimum of 3 months to minimize memory effects. This allowed us to determine the reproducibilities of the radioligists' region selections.

A comparison of the ROIs selected by each of the 5 radiologist readers (a-e) and the TROIs for an image containing 3 TROIs is shown in Fig. 1, below. The CAD selected ROIs are also compared with the TROIs in part (f) of this figure.

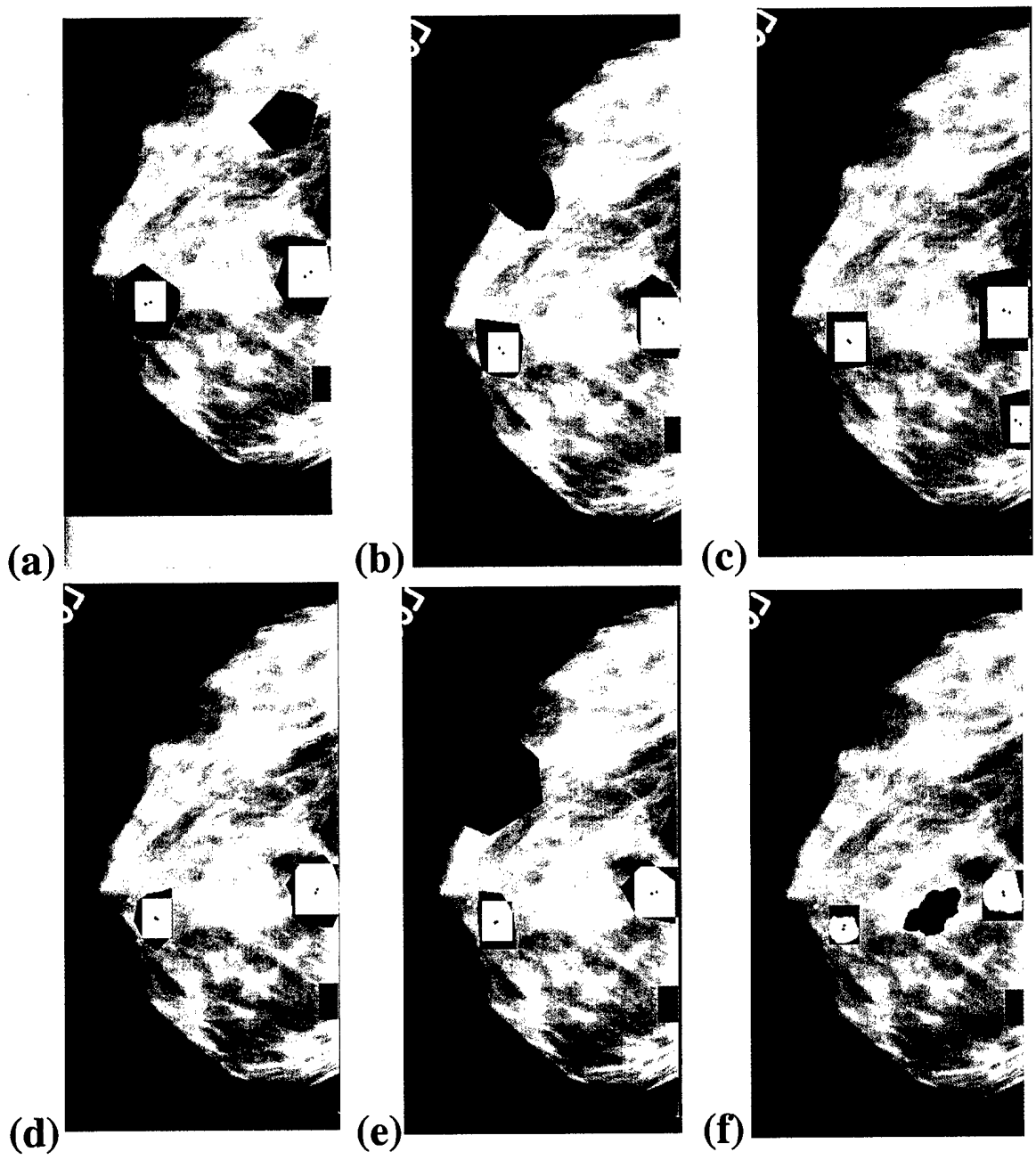


FIG. 1. Comparisons of ROIs selected by radiologists with true mass regions (a-e) and the CAD program with true mass region (f). The radiologist and CAD selected regions are filled-in in red, and the true regions are filled-in in green. The intersections are displayed in yellow. Note that for this particular mammogram, radiologist (d) chose to select 2 regions instead of 3. Also, radiologist (c) was the only one whose selected regions intersected all 3 TROIs. [** Unpublished Data **]

The average values of the largest overlap indices between the radiologist selected ROIs and the CAD selected ROIs are listed for each radiologist for each reading session in Table I, below.

BEST AVAILABLE COPY

Table I. Overlap indices between the ROIs selected by the radiologists and by the CAD computer program. The averages of the largest overlap indices for each image are listed for each radiologist at each reading session. These are averages for 200 images and up to 3 ROIs per image. [** Unpublished Data **]

	Radiologist a	Radiologist b	Radiologist c	Radiologist d	Radiologist e
Reading #1	94%	84%	51%	45%	82%
Reading #2	98%	86%	47%	43%	65%

The agreement between the radiologist selected ROIs for spot imaging and the CAD selected ROIs had a wide range (43% to 98%). The primary reason for this is that the probability of overlap is highly dependent upon the numbers and sizes of the ROIs selected by the radiologist, with greater numbers and ROI sizes corresponding with a greater likelihood that there would be more overlap between the radiologist and CAD selected ROIs. Although each radiologist could select up to 3 ROIs in each image, we found there was a large range in the actual number of ROIs selected by the various radiologists. (The mean number of ROIs/image ranged from 0.66 for one of the radiologists to 2.94 for another.) In analyzing the data, we found there was a very high correlation of 0.99 between the overlap indices in Table I (above) and the average number of ROIs selected per image by the 5 radiologist observers.

The reproducibility of the overlap indices between the radiologist and CAD selected ROIs is very good (within 4% for 3 radiologists and 17% for one) indicating the radiologists are consistent in their selection of ROIs.

The average overlap indices between the TROIs and the ROIs selected by the radiologists and the CAD readers are listed in Table II, below.

TABLE II. Overlap indices between the ROIs selected by the readers and the True Regions of Interest.(TROIs). The averages of the largest overlap indices for each TROI in each image are listed, as well as the averages for all TROIs. The overlap index in this case is defined to be the area of the intersection divided by the smaller of the TROI and reader ROI areas, in percent. There were 83 images with one TROI, 13 with two TROIs and 2 with three TROIs.
[** Unpublished Data **]

	TROI #1 n=83	TROI #2 n=13	TROI #3 n=2	All TROI n=115
Radiologist a				
Reading #1	88%	58%	50%	84%
Reading #2	94%	74%	46%	90%
Radiologist b				
Reading #1	88%	70%	49%	85%
Reading #2	88%	66%	46%	84%
Radiologist c				
Reading #1	70%	42%	47%	66%
Reading #2	71%	44%	44%	67%
Radiologist d				
Reading #1	69%	45%	0%	65%
Reading #2	66%	48%	50%	64%
Radiologist e				
Reading #1	87%	50%	37%	81%
Reading #2	85%	67%	47%	82%
Computer (CAD)	78%	53%	0%	73%

The agreement between the radiologist selected ROIs and the true mass regions (TROIs) is good. For the entire set of 115 TROIs (Table II, column 4), the overall average overlap indices for the radiologists' ROI vs. TROI comparisons ranged from 64% to 90% with a mean value for all 5 radiologists of 76.8% +/- 10-%, and the overall average overlap index for the CAD ROI vs TROI comparisons was 73%. The radiologists were very consistent in selecting similar regions at each reading session, with the overlap indices with all TROIs for four of the five radiologists being reproducible to within 1%. (Table II, column 5).

The average percentages of hits between each reader's selected ROIs and the TROIs are listed in Table III, below.

TABLE III Average percentages of "hits". A hit is defined to occur whenever the overlap index between the reader ROI and the TROI is greater than or equal to 25%.
Values are listed for each reading session. [** Unpublished Data **]

	Reading #1	Reading #2
Radiologist a	89%	93%
Radiologist b	90%	89%
Radiologist c	70%	70%
Radiologist d	67%	69%
Radiologist e	87%	87%
CAD	80%	

The percentages of "hits" for which the overlap indices were greater than or equal to 25% for all TROIs were very good with a range of 67% to 93% for the radiologist vs TROI comparisons and 80% for the CAD vs TROI comparison The reproducibilities of the hit results for the radiologists were within 4% for the two reading sessions. The mean percentage of hits for all five radiologists was 81% +/- 11.1% for the first reading and 82% +/- 11.3% for the second reading. The performance of the CAD method (80%) is therefore very close to the average of the experienced radiologists.

The good agreement between the CAD selected ROIs and the TROIs indicate the CAD routines have promise in an automated implementation of stereo spot mammographic imaging. The results of this observer study were presented at the 2003 AAPM meeting in San Diego, CA and **a detailed manuscript describing this study has been submitted to the journal Medical Physics.[1]**

B) Task 2: Develop Secondary Collimator

As reported last year, the manufacturing of the secondary "spot" collimator was finished during the third year of this project. This year we analyzed the collimator blade positioning requirements for stereo spot image acquisition. The geometric considerations for this task are illustrated in Fig. 2, below [** Unpublished Data **]

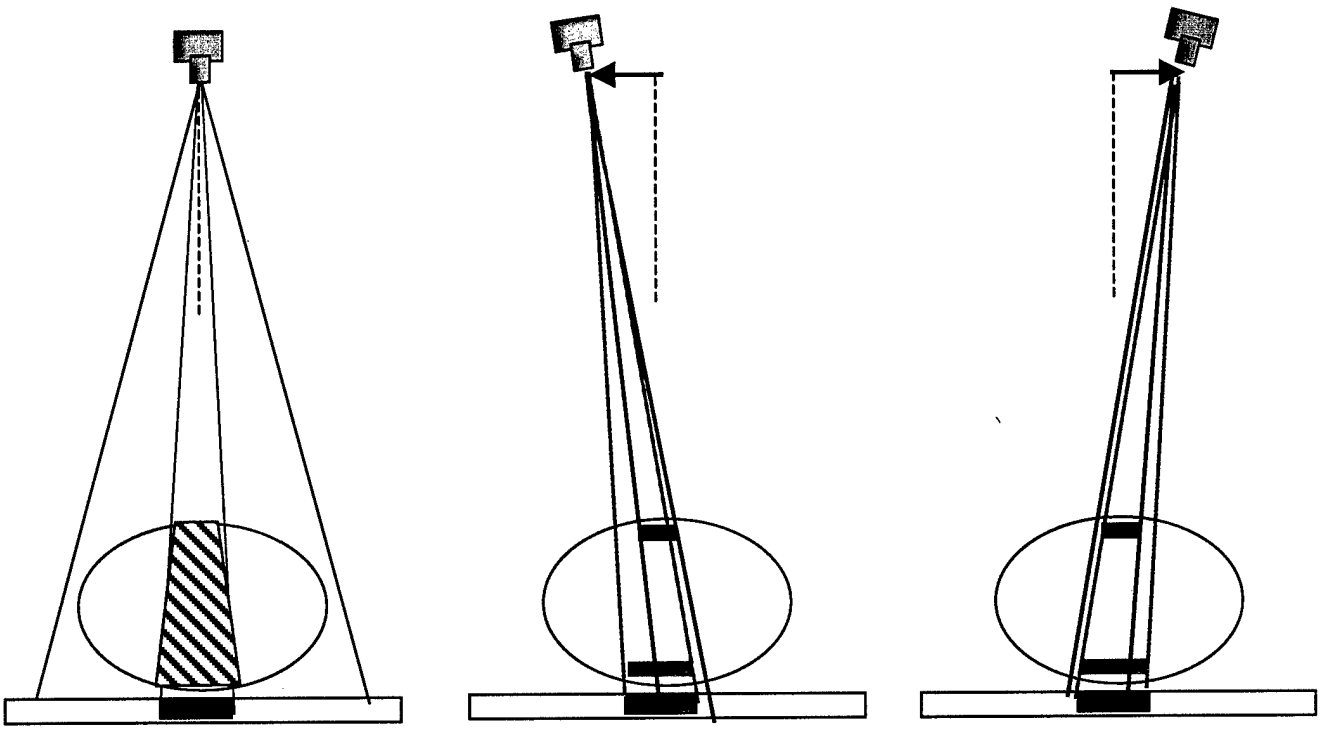


FIG. 2. Geometrical considerations for stereo spot imaging. (a) Full-field (0°) geometry showing CAD determined spot region in purple and corresponding lesion region in the breast as a crosshatched area. (b) collimated beam boundaries required to image lesion if it is located near top (blue) and bottom (orange) of crosshatched region for left-eye stereo spot image, (c) corresponding beam boundaries for right-eye stereo spot image.

The CAD ROI that eventually will be used for spot collimation is determined from the single projection full field digital mammogram. As shown in Fig. 2 (a) , the lesion associated with the ROI can be located anywhere within the shaded-in trapezoidal shaped region of the breast. The collimator blade positions that are required to insure full coverage of the lesion for the stereo spot imaging are highly dependent upon the height of the lesion. This is illustrated for two extreme positions of the lesion (near the top of the breast (blue) and near the bottom of the breast (orange) in Fig. 2 (b) for the acquisition of the left-eye stereo image, and Fig. 2(c) for the acquisition of the right-eye stereo image. To insure that the entire ROI is contained within the projected areas, for the left eye stereo spot image (Fig. 2 (b)), the left collimator blade should restrict the beam to strike the detector at the left edge of the CAD ROI that was determined in the full-field digital mammography image (shown in purple in Fig. 2(b)), and the right collimator blade should restrict the beam along the rightmost blue line extending from the focal spot through the right most point on the upper blue lesion position. For the right-eye stereo spot image (Fig. 2 (c)) the right collimator blade should restrict the beam to strike the detector at the right edge of the CAD ROI that was determined in the full-field digital mammography image (shown in purple in Fig. 2(c)), and the left collimator blade should restrict the beam along the leftmost blue line extending from the focal spot through the left most point on the upper blue lesion position. The relationship between the collimator blade position relative to the projected collimator edge shadow location on the detector was derived for the 0 degree (full-field) and the 6-degree left-eye and 6 degree right-eye shifted x-ray tube positions from digital mammography images taken in each orientation with the blades moved to locations that were known amounts of steps from the collimator blade/stepper motor homing positions.

To illustrate how the shadow positions on the detector are determined, consider the situation shown in Fig. 3, below for the GE full-field digital mammography system and a compressed 7-cm thick breast. The source to image distance for the GE system is 66 cm, and the fulcrum point for stereo shifting is 20 cm above the detector, so ignoring the small arc for the tube shift, a 6-degree tube shift corresponds with a lateral translation of about 4.8 cm. (= (66 cm-20 cm) x tan 6^{o}) Assume the maximum height of the lesion is 6 cm above the digital detector. If the rightmost limit of the CAD determined ROI for the lesion in the 0 degree (full-field digital mammography) image is at a position (5 cm, 0 cm), then the corresponding x-position within the breast at a height of 6 cm can be determined from the equation for the line passing from the point (0, 66) to (5, 0). It is found to be 4.54 cm The coordinates of the point in the breast (xb, 6) is thus (4.54, 6). The desired intersection at the detector (x, 0) is computed by extrapolation along the line connecting the shifted x-ray tube focal spot position (-4.8, 66) with this (4.54, 6) point in the breast. The computed intersection is at the point (5.48, 0), which is input into the above mentioned collimator blade calibration relationship to determine the collimator blade position. The thickness of the compressed breast of a given patient is measured by the compression paddle of the mammography system and can be input into the collimator driving software for patient-specific spot imaging.

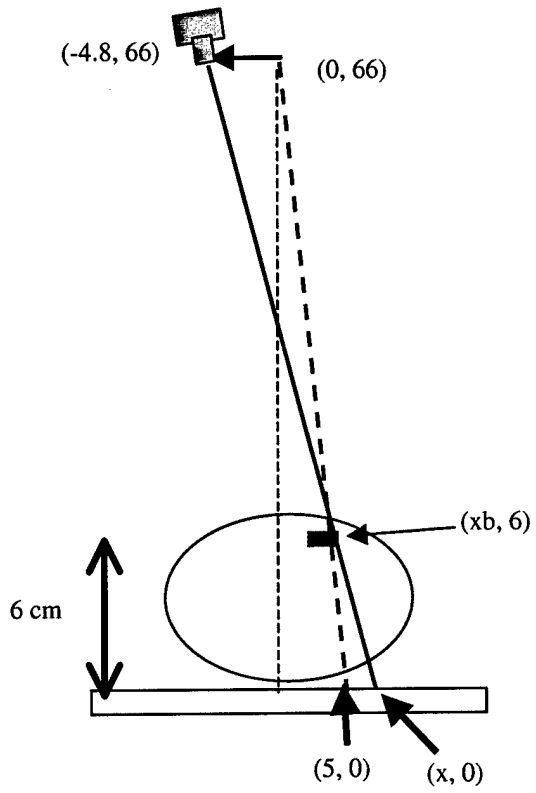


FIG. 3 Example of geometry employed to determine projected position of right spot collimator blade for GE full-field digital mammography system. [** Unpublished Data **]

C) Task 3: Develop system to restrain breast during changeover from full-field to spot compression paddle and reduce tension as spot paddle presses against breast.

A prototype system for accomplishing this task was designed and built during the first year of the project . Since we changed our approach as described in the Modified Statement of Work (dated January 7, 2002) , such a device is not necessary for the new stereo spot technique that we are pursuing. Therefore, there was no further development of this device during the fourth year of the project.

D) Task 4: Develop system to position spot compression paddle

A prototype spot paddle positioner was designed and built during the first year of the project. This device is also not needed for the new stereo spot technique that we are pursuing. Therefore, there was no further development of this device during the fourth year of the project.

E) Task 5: Evaluate Automated Exposure Optimization (AOP) techniques for spot imaging application on commercial full-field digital mammography device.

Because we did not receive the modular breast phantom that we planned to use for this task until late in July and because this phantom did not meet our design specifications, we were unable to compete this task . This is one of the reasons we applied for an additional one-year no cost extension of this project. This extension was approved, and we plan to perform the AOP stereo spot studies during the fourth year of our project.

F) Task 6: Develop breast phantoms

We submitted our original designs for the breast phantom for stereo spot mammography research to the manufacturer in November of 2002. The phantom was to be modular with sections containing simulated masses and microcalcifications that could be positioned at different depths, and other sections containing simulated glandular tissue that would overlap the masses and microcalcifications. The phantom was supposed to produce images with a patient-like mammographic appearance. All of the requirements are very important for our work, because they allows us to test the ability of stereo imaging to reduce the camouflaging effect of overlying and underlying tissues. The manufacturer found it difficult to meet our requirements and this resulted in an unexpected manufacturing delay. We received the phantom on July 29, 2003, which was over 9 months after our original order. Unfortunately, when we imaged the phantom, we found that the manufacturer had not met our design goals. (See images in Fig. 4, below.) The glandular tissue regions are located too close to the outer edge of the phantom and have little or no overlap with the simulated masses, and the glandular tissues have a swirled appearance with sharp edges that is very different from the appearance in real mammograms. We have contacted the manufacturer and requested that they make changes to better meet our design goals. They have agreed to do this within a month or two.

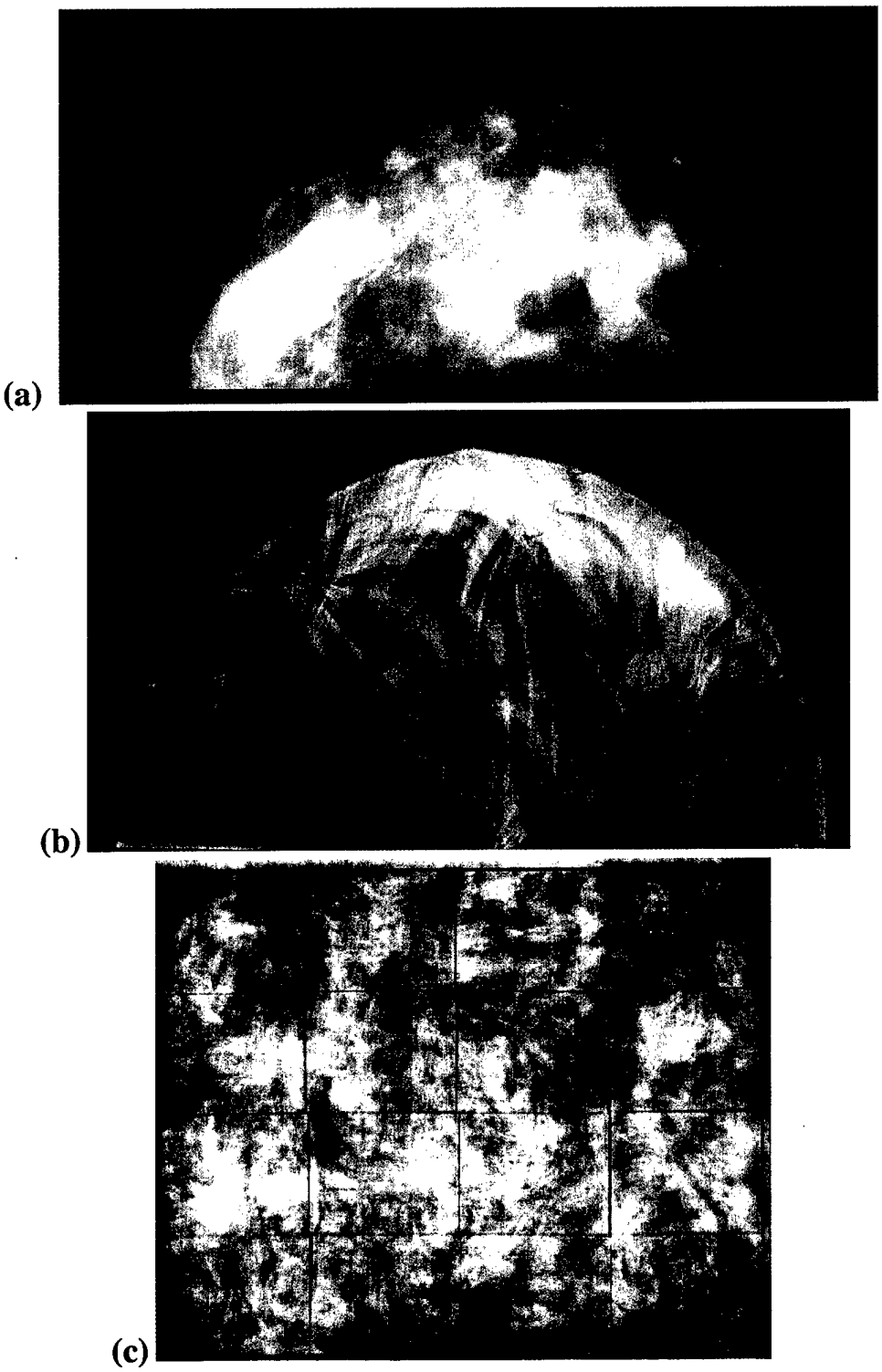


FIG. 4 (a) Desired mammogram appearance of phantom from our original design specifications, (b) Mammogram of the current phantom (c) X-ray image of a different phantom built by the same manufacturer for another client. Our original design specification called for an image texture similar to the texture in this image, which is also similar to that of the patient mammogram in (a). This mammogram (c) was published in reference 1.
[** Parts (a) and (b) Unpublished Data **]

Since we still do not have an adequate phantom, we have been unable to complete tasks requiring the phantom including Task 5 (Evaluation of Automated Exposure Optimization for

stereo spot imaging), and Task 9 (Develop and implement changes in automated spot collimator design for prototype spot stereo mammography image acquisition system.)

Once an adequate phantom is obtained, we will be able to perform additional investigations comparing contact stereomammography with magnification stereomammography (Task 8) and preliminary investigations comparing stereospot mammography with tomosynthesis. Currently, there is great interest in the development of tomosynthesis techniques for imaging the breast in 3-D. [2] The modular property of the breast phantom makes it ideal for the comparison of stereo spot imaging with tomosynthesis. Although we did not propose to perform such a comparison in our original statement of work, an additional no-cost time extension of 1 year would provide us with the opportunity to perform this comparison. We anticipate that this study could lead to future areas of 3-D breast imaging research.

G) Task 7: Explore possible advantages of using stereo-spot mammography instead of single-projection, spot compression with spot collimation mammography for imaging overlapping structures

This task was completed during the second year of the project, and reported on in the second year annual report. We performed further analysis of the results and presented them during the third year of the project at the American Association of Physicists in Medicine (AAPM) meeting in Montreal, Canada, July 14-18, 2002,[3] and at the ERA of Hope meeting in Orlando, Florida, September 25-28, 2002. [4,5] We will perform additional studies related to this task during the next year of this project when we receive a modular breast phantom that satisfies our design requirements.

H) Task 8: Compare Contact and magnification stereomammography

We finished this task during the third year of the project. We wrote a paper on our investigation of depth measurements made with 3-D stereo cursors. That paper, entitled "The effects of stereo shift angle, geometric magnification and display zoom on depth measurements in digital stereomammography, " was published in the Medical Physics journal after our last report (Volume 29 (11), November 2002, pages 2725-2734) [6] We also presented the results of our study comparing the ability of observers to distinguish positions of simulated fibrils at different radiation doses, stereo angles, and geometric magnifications in a poster at the SPIE-Medical Imaging meeting at San Diego, CA in February, 2002) [7] and at the ERA of Hope Meeting [8]. This year, we submitted a paper entitled, "Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study," to Physics in Medicine and Biology 2003 [9]. This paper has been accepted for publication.

I) Task 9: Develop and implement changes in automated spot collimator design for prototype spot stereo mammography image acquisition system.

We calibrated the secondary collimator for single projection cranial caudal (CC) imaging last year and acquired stereo projection images with the spot collimator. When we receive the corrected modular breast phantom, we plan to continue our studies, more fully testing the geometric relations described above in Task 2.

J) Task 10: Design second generation auto stereo spot mammography system

This final task will be performed during the next year of the project after we have gained more experimental experience with our first generation auto stereo spot mammography system.

(6) Key Research Accomplishments

- Developed software to display true regions of interest (TROIs) corresponding to true masses on the digitized mammograms used in the observer study.
- Developed overlap indices for comparing degree of agreement between radiologist selected regions of interest (ROIs) and CAD selected regions of interest as well as between radiologist selected ROIs and TROIs and between CAD selected ROIs and TROIs.
- Completed radiologist observer study using program that displays images at higher (200 micron and 400 micron pixel) resolution and analyzed overlap indices. Found good agreement between CAD selected ROIs and TROIs indicating CAD has promise for selecting suspicious regions in the implementation of an automated stereo spot imaging technique.
- Analyzed geometric considerations and performed calibration for secondary collimator to insure collimator blades are opened sufficiently to guarantee complete coverage of the suspicious region in stereo spot image acquisition.
- Obtained modular breast phantom for stereo spot imaging experiments. Imaged phantom and found it did not satisfy our design requirements. Discussed this with manufacturer and they agreed to produce a corrected phantom for us in the near future.

(7) Reportable Outcomes

Manuscripts

1) MM Goodsitt, HP Chan, KL Darner and LM. Hadjiiski. The Effects of Stereo Shift Angle, Geometric Magnification, and Display Zoom on Depth Measurements in Digital Stereomammography Medical Physics 2002; 29(11); 2725-2734.

2) HP Chan, MM Goodsitt, LM Hadjiiski , JE Bailey, K Klein , KL Darner, B Sahiner. Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study. Accepted for Publication in Physics in Medicine and Biology, 2003

3) J Wei, HP Chan, M.A. Helvie, M.A. Roubidoux, B Sahiner, L.M. Hadjiiski, C. Zhou, S. Paquerault, T. Chenevert, M M. Goodsitt, Correlation between Mammographic Density and Volumetric Fibroglandular Tissue Estimated on Breast MR Images, Accepted for publication, pending revision, in Medical Physics, 2003

Abstracts and Presentations

1) Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Helvie MA, Bailey JE, Roubidoux M, Paramagul C, Sahiner B, Petrick N. Automated Stereo Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists and a Computer. Presented at the 45th Annual Meeting of the American Association of Physicist in Medicine (AAPM) in San Diego, CA August 10-14, 2003. (Medical Physics 2003;30:1456)

2) Chan HP, Goodsitt MM, Hadjiiski L, Roubidoux MA, Bailey JE, Helvie MA, Lydick JT, Sahiner B. ROC study comparing radiologists' performances in evaluating breast lesions on stereoscopic and single-projection digital specimen mammograms. Presented at the 45th Annual Meeting of the American Association of Physicists in Medicine (AAPM) in San Diego, CA August 10-14, 2003. (Medical Physics 2003;30:1456)

CONCLUSIONS

During the fourth year of this project, we have pursued developing the automated stereo spot imaging technique for improved visualization of dense regions in mammograms. This is the technique that we concluded during our second year would be more practical and offer superior visualization of superimposed structures than our originally planned method of combining spot automated spot compression with automated spot collimation. We finished our observer studies comparing the suspicious regions that different radiologists would choose for spot imaging with those selected by a CAD computer program. We found there was significant variation in the overlap between the radiologist and CAD selected ROIs due mostly to variations in the numbers of radiologist selected ROIs per image and the sizes of the radiologist selected ROIs, with greater numbers and larger sizes having greater probabilities of overlap. We also performed a study comparing the radiologist and CAD selected regions to true mass regions determined by an experienced MQSA certified radiologist using mammograms, biopsy images and pathology results. The CAD-selected ROIs had a 73% overlap with the TROIs, which was comparable with the average overlap result of 76.8% +/- 10% for the 5 experienced radiologists who participated in our observer study. Thus, CAD determined ROIs could potentially be useful for a screening technique that includes stereo spot mammography. We determined the geometric requirements for the automated collimator to insure the suspicious region in a full-field digital mammogram is adequately covered during stereo spot image acquisition. We obtained our modular breast phantom containing simulated overlapping masses, dense regions and microcalcifications at the very end of our fourth year. Our initial images of this phantom showed that it did not satisfy our design requirements. In particular, the dense regions did not

have sufficient overlap with the simulated masses, and the phantom texture did not of simulate that of a real mammogram. We therefore were not able to complete our planned phantom experiments. The manufacturer has promised to correct the problems and produce an improved modular breast phantom for us in the near future. We plan to perform the phantom studies during the additional no-cost time extension year that was recently approved for this project, and we plan to perform a preliminary investigation of spot tomosynthesis, which is another very promising 3D x-ray imaging technique.

So What

A key limitation of conventional x-ray mammography is the inability to optimally image regions of dense and overlapping tissue. The new full-field digital mammography systems reduce but not entirely eliminate this problem. Additional views including spot imaging are still employed with the digital systems to better analyze suspicious tissue regions. The system we are developing adds 3 features that will further optimize spot imaging of suspicious tissue. First computer vision techniques will be employed to automatically find suspicious regions that warrant spot imaging. Second, the system will automatically collimate the x-ray beam to the suspicious region. This will reduce the amount of x-ray scatter that strikes the detector thereby reducing image noise. It should also restrict the region in which the automated optimization of parameters (AOP (kVp, target, and filter)) and phototiming are determined, which may improve the penetration of the spot region. Third, the spot image will be acquired using a stereoscopic method. This will produce a 3-D image allowing the radiologist to see the depth separation between lesions and overlapping tissues and allowing the radiologist to see through any "cloud" of dense tissue. Stereoscopic imaging also permits improved visualization of the locations and 3D shapes of lesions and microcalcification clusters, and of lesion borders, all of which are important to the radiologists in making their diagnoses. Finally, the automated spot collimation stereo mammography technique is performed with the breast in the same location as in the full-field mammogram, eliminating the guesswork associated with breast re-positioning in conventional spot compression mammography. Thus, in theory, the automated spot collimation, stereomammography technique should be of great benefit. The challenge will be to develop a convenient and practical system that could be employed in screening, which is the long-term goal of our project

(9) REFERENCES

1) MM Goodsitt, HP Chan, JT Lydick, CR Gandra, NG Chen, MA Helvie, J Bailey, MA Roubidoux, C Paramagul, CE Blane, B Sahiner, NA Petrick, Automated stereo spot mammography: A Comparison of spot imaging regions selected by radiologists and a computer. **Submitted to Medical Physics, 2003**.

2) Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH , Landberg CE, Slanetz PJ, Giardino AA, Moore R, Albagl Di D, DeJuleMC, Fitzgerald PF, , et al,. "Digital tomosynthesis in breast imaging," Radiology 1997; 205: 399-406

3) Goodsitt M, Chan H-P, Gandra C, Chen N, Helvie M, Klein K, Bailey J, Paramagul C. Automated Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists. Presented at the 44th Annual Meeting of the American Association of Physicists in Medicine in Montreal, Quebec, Canada, July 14-18, 2002.(Medical Physics 2002;29:1307)

4) Goodsitt MM, Chan HP, Gandra CR, Chen NG, Helive MA. Automated stereo spot mammography for improved imaging of dense breasts. Poster presentation at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, Sept 25-28, 2002. P28-11 in Proceedings, Volume II.

5) Goodsitt MM. Automated Stereo Spot Mammography for Improved Imaging of Dense Breasts. Presented in a Symposium entitled Digital Imaging: Diagnostic Potential and Enhancing Availability at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, September 25-28, 2002. Page 30 of Program.

6) Goodsitt MM, Chan HP, Darner KL and Hadjiiski LM. The Effects of Stereo Shift Angle, Geometric Magnification, and Display Zoom on Depth Measurements in Digital Stereomammography Medical. Physics 2002; 29(11): 2725-2734)

7) Chan HP, Goodsitt MM, Hadjiiski L, Bailey JE, Klein K, Darner KL, Paramagul C. Digital stereomammography: observer performance study of the effects of magnification and zooming on depth perception. Proc SPIE 4682. 2002: 163-166.

8) Chan HP, Goodsitt M, Hadjiiski L, Helvie M, Bailey J, Klein K, Roubidoux M. Development of digital stereo imaging technique for mammography. Poster presentation at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, Sept 25-28, 2002. P28-6 in Proceedings, Volume II.

9) Chan, HP, Goodsitt MM, Hadjiiski LM, Bailey JE, Klein K, Darner KL, Sahiner B. Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study. Accepted for publication in Physics in Medicine and Biology 2003

(10) Appendix

The following publications in the current year as a result of this grant are enclosed with this report.

1) MM Goodsitt, HP Chan, KL Darner and LM. Hadjiiski. The Effects of Stereo Shift Angle, Geometric Magnification, and Display Zoom on Depth Measurements in Digital Stereomammography Medical Physics 2002; 29(11); 2725-2734.

2) HP Chan, MM Goodsitt, LM Hadjiiski, JE Bailey, K Klein, KL Darner, B Sahiner. Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study. Accepted for Publication in Physics in Medicine and Biology, 2003 (Galley Proof)

3) Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Helvie MA, Bailey JE, Roubidoux M, Paramagul C, Sahiner B, Petrick N. Automated Stereo Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists and a Computer. Presented at the 45th Annual Meeting of the American Association of Physicist in Medicine (AAPM) in San Diego, CA August 10-14, 2003. Abstract (Medical Physics 2003;30:1456)

The effects of stereo shift angle, geometric magnification and display zoom on depth measurements in digital stereomammography

Mitchell M. Goodsitt,[a] Heang-Ping Chan, Katie L. Darner, and Lubomir M. Hadjiiski
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-0030

(Received 23 January 2002; accepted for publication 15 August 2002; published 28 October 2002)

We are developing virtual three-dimensional (3-D) cursors for measuring depths in digital stereomammograms. We performed a study to investigate the effects of stereo shift angle, geometric magnification, and display zoom on the accuracy of depth measurements made with a virtual 3-D cursor. A phantom containing 50 low contrast fibrils at depths ranging from 1 to 11 mm was imaged with a full-field digital mammography system. Left- and right-eye images were generated at stereo shift angles of ±3° and ±6°, using either contact or 1.8× geometric magnification geometry. The images were viewed on a high-resolution stereoscopic display system in normal and 2× zoom mode. Observers viewed the images with stereo glasses and adjusted the depth of a cross-shaped virtual cursor to best match the perceived depth of each fibril. The results for two trained observers with excellent stereo acuity were nearly identical when viewing the same images. The average root mean square errors for the two observers were 1.2 mm (±3° contact, no zoom), 1.3 mm (±3° contact zoom), 0.8 mm (±6° contact, no zoom), 0.6 mm (±6° contact, zoom), 0.8 mm (±3 magnification, no zoom), 0.7 mm (±3° magnification, zoom), and 0.2 mm (±6° magnification, no zoom). One observer repeated the entire study for two additional fibril phantom configurations. Combining all the results, we found that for the contact geometry increasing the stereo shift angle from ±3° to ±6° improved the depth measurement accuracy by factors of about 1.2–4.0. Zooming did not provide observable improvement in the depth measurement accuracy; sometimes having no effect, sometimes improving the accuracy, and other times reducing the accuracy, with no general trends. Its effect is likely within experimental errors. However, the stereo effect was more readily visualized in the zoom mode. Geometric magnification improved the depth measurement accuracy. The best accuracy among all cases was about 0.2 mm, obtained with geometric magnification using a stereo angle of ±6°. This is the mode we recommend for obtaining accurate depth measurements with virtual cursors in stereomammograms. © *2002 American Association of Physicists in Medicine.* [DOI: 10.1118/1.1517615]

Key words: stereomammography, stereoscopic, virtual cursor, 3-D imaging

I. INTRODUCTION

Tissue superposition makes it difficult to accurately interpret conventional mammograms. Such mammograms are acquired using a single projection method whereby (ignoring scatter) the density at a point in the image represents the summation of the attenuation of all tissues along a ray extending from the x-ray tube focal spot to that point. The superpositon of tissues along the rays decreases image contrast and can result in the camouflaging of masses and microcalcifications within dense tissue. It can also lead to superimposed structures having mass-like appearances. The superpositon problem can be reduced or eliminated by generating and viewing 3-D mammograms via multiprojection techniques such as stereoradiography,[1–7] tomosynthesis,[8–10] and computed tomography.[11,12]

We have been investigating digital stereomammography. This is a computerized version of an analog technique that was first described by Warren in 1930.[13] Both the new and old techniques involve taking two separate mammograms, one with the x-ray tube at a positive angle (e.g., +3°) relative to a normal to the detector and the other with the x-ray tube at an equal but opposite angle (e.g., −3°) about the normal. One of the images is viewed with the left eye and the other with the right eye. Our brain fuses the images together to create a 3-D effect. The old technique required taking two films in roughly the same projection. As such, it had several disadvantages, including at least twice the x-ray dose, film cost, and processing time. It also required increased procedure time and radiologist viewing time. Radiologists, in general, eventually decided that these disadvantages outweighed the 3-D visualization advantage, and film stereomammography was discontinued. Digital mammography eliminates or reduces most of these advantages, thereby making digital stereomammography a potentially viable technique. In contrast to screen-film systems, which have sigmoid-shaped response curves, digital detectors have a linear response. Thus, the response curve of the digital detector does not degrade image contrast at lower doses, and it may be possible to utilize half the normal dose for each digital image. The two images of the stereo image pair will be integrated by the obsever's eye–brain system to yield about the same signal-to-noise ratio as in a single image taken with the same total dose.[14] With digital systems, image processing and display are almost instantaneous. Also, the method of examining the

images, displayed on a television monitor and viewed with liquid crystal display (LCD) glasses that are synchronized so the left eye sees one image and the right the other, is more convenient and less time consuming than the film counterpart.

We have been investigating the use of virtual 3-D cursors for measuring depths in digital stereomammograms.[3,5,6] Leduc *et al.*[1] have also performed research in this area. The 3-D cursors are generated with computer graphics and are overlaid on the digital mammograms. In our initial studies reported previously,[3] we determined the accuracies of observers' measurements of the depths of horizontally and vertically oriented nylon filaments that simulate fibrils in mammograms. We found that when observers used a cross-shaped cursor, they could determine depths of vertically oriented fibrils with accuracies [root mean square (rms) errors] of 0.4–1.3 mm, but their accuracies were degraded for horizontally oriented fibrils (rms errors of 1.9–4.2 mm). In a subsequent study,[5] we found that use of a comb-shaped cursor improved the accuracies (reduced the rms errors) of observers' depth measurements of the horizontally oriented fibrils by 0.1–1.4 mm. With this cursor, two of the observers were able to measure the absolute depths of the horizontal fibrils with much improved accuracies of 0.8–1.0 mm. The images for our previous studies were generated with a Fischer (Denver, CO) MammoVision Sterotaxic unit, using a stereo shift of ±2.5°. More recently, our Radiology department obtained a GE (Milwaukee, WI) Senographe 2000D full-field digital mammography system. In this report, we describe a study that was performed using images acquired with this system. We investigated the effects to stereo shift angle, geometric magnification, and display zoom on the accuracy of depth measurements made with a virtual 3-D cursor.

II. MATERIALS AND METHODS

A. Phantom

We employed the same multi-layered fibril phantom that was used in our previous studies.[3,5,6] This phantom consists of six 1 mm thick Lexan sheets each separated by 1 mm spacers. A 5×5 matrix of 8 mm long, 0.53 mm diameter nylon fibrils is placed on the plates with 25 fibrils oriented vertically (perpendicular) and 25 horizontally (parallel) relative to the stereo shift direction. The depths and orientations of the fibrils were randomized and organized such that one horizontal fibril crossed one vertical fibril at each of the 25 matrix positions. The order of the Lexan layers could be changed to create many independent phantom configurations. For our present experiments, we rotated the phantom 45°. This resulted in 25 of the fibrils being oriented at +45° and 25 at −45° relative to the stereo shift direction. Thus, all fibrils had both horizontal and vertical components.

B. Stereo image acquisition

As mentioned above, the images were generated with a GE Senographe 2000D digital mammography system. The pixel size for this system is 100 μ in contrast to the 50 μ pixel size of our previous images obtained with the Fischer MammoVision stereotaxic unit. The GE system employs a digital detector consisting of a CsI:Tl light converter and an *a*-Si active matrix flat panel unit with photodiodes and TFT pixel switches.[15] The detector measures 23 cm×19 cm. Stereo images are traditionally produced by shifting the x-ray tube to the right and left of an axis perpendicular to the detector. Usually, the total x-ray tube shift is 10% of the focus-to-detector distance, which corresponds, with a stereo shift angle of about ±3° [$=\tan^{-1}(0.1/2)$]. For phantom imaging, an alternative to shifting the x-ray tube is shifting the phantom. This is illustrated in Fig. 1. We employed this method for all of the contact images in this study. We built a phantom sliding device with positioning marks at the locations for the two shift angles that were studied (±3° and ±6°). Note for illustration purposes, to simplify the comparison of the methods, that the x-ray tube in Fig. 1 is shown rotating about a fulcrum at the center of the detector, with a stereo shift angle, θ. In the calculation of the shift distance, w, for our phantom-shift imaging setup, the actual geometry was used (i.e., a source-to-detector distance of 66 cm, a fulcrum 46 cm from the focal spot, and a magnification-stand-to-detector distance of 26.4 cm). The small difference in the focus-to-detector and focus-to-object distances for the tube-shift and phantom-shift methods, when the tube shift involves x-ray tube rotation about a fulcrum, was neglected in our study. For the actual contact and magnification geometries that were employed, it can be shown that these differ-

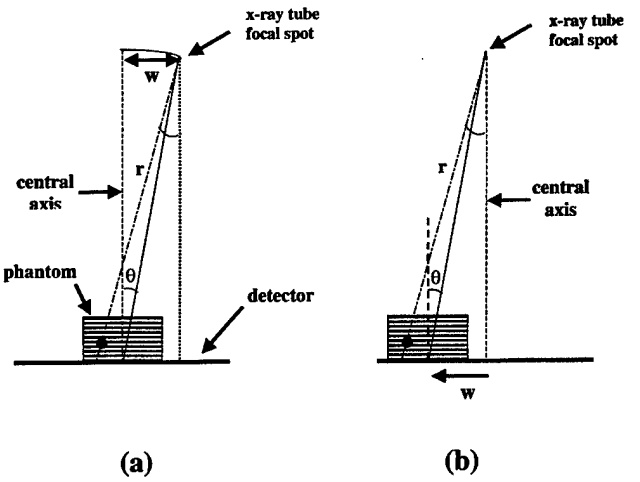


FIG. 1. Tube shift (a) and phantom shift (b) methods of stereo image acquisition. Both sketches illustrate the geometries for the generation of the right-eye images. In (a) the x-ray tube focal spot shifts a distance w to the right. In (b), the x-ray tube focal spot remains fixed along the central axis of the detector, and the phantom shifts a distance w to the left. The sketches illustrate the equivalency of the geometries of the two methods. Notice that they ray r passes through the black circular object in the phantom at the same angle in both cases. The corresponding drawing for the left-eye images of the stereo pairs would have the x-ray tube in (a) shifted to the left a distance w, and the phantom in (b) shifted a distance w to the right.

ences have a 0.6% or smaller effect on the accuracy of the results, which is essentially negligible (e.g., 0.6% of 1.0 mm accuracy=0.006 mm).

When using the tube-shift and phantom-shift methods, it is desirable to align the resulting images such that an object (e.g., a fibril) in contact with the bottom of the phantom does not shift. By doing so, all depths or distances of objects in the phantom will be measured relative to the back surface of the phantom. We achieved the desired zero shift by placing a fiducial marker on the top surface of the slider on which the phantom was placed and digitally translating the resulting left- and right-eye images so that the fiducial markers coincided. For magnification mammography, especially at larger stereo shift angles, the phantom shift method cannot be used because only a portion of the phantom will project to within the field of view of the detector due to the limited size of the detector. For our magnification techniques, we employed the phantom-shift method for the $\pm 3°$ stereo image acquisition and the tube-shift method for the $\pm 6°$ stereo image acquisition. All images in this study were obtained using a technique of 30 kVp, Rh filter, Rh target, 63 mAs. The large (nominal 0.3 mm) focal spot was employed for the contact images and the small (nominal 0.15 mm) focal spot for the ($1.8\times$) geometric magnification images. The scatter-rejection grid was removed for magnification image acquisition.

C. Stereo image display

The stereoscopic display system that was employed for this study consisted of a Barco-Metheus (Beaverton, OR) model 1760S stereo graphics board in a SUN Microsystems (Palo Alto, CA) Ultra 10 computer. The Metheus board operates in a page flipping stereoscopic mode whereby the left- and right-eye images are displayed sequentially, one after the other. This board is capable of displaying $1408\times1408\times8$ bit progressive-scan images at a refresh rate of 114 Hz. The images in our study were displayed on a 21 in. Barco model 521 monitor and viewed with NuVision (Beaverton, OR) LCD stereoscopic glasses. We employed in-house developed software to display, pan, zoom, and adjust the contrast and brightness of the images.

D. Virtual cursor

We developed software to generate the virtual cursors and display their x, y, and z positions. The 3-D nature of the cursor is achieved by introducing offsets in the horizontal positions of the representations of the cursor in the left- and right-eye images. The z coordinate is equal to the offset. When the offset is 0, the cursor is at the same x, y position in both images, and it appears stereoscopically to be at the depth of the monitor screen. As the horizontal offset between the cursor positions is increased in one direction (e.g., left), the cursor appears to move closer to the observer, and as the offset is increased in the opposite direction (e.g., right), the cursor appears to move toward or into the monitor.

E. *Z*-coordinate calibration

The z coordinate was calibrated by imaging thin wires placed on the steps of a solid acrylic step wedge accurately milled with known step heights. The step wedge was imaged with the GE digital mammography system using the same phantom shift or x-ray tube shift as for the images of the fibril phantoms under the corresponding imaging conditions, and using the fiducial marker alignment technique described previously. The thin wires that were employed for calibration were oriented perpendicular to the tube shift direction. The resulting stereoscopic images were viewed without the stereo glasses and the left- and right-eye cursor positions were adjusted to overlay the left- and right-eye images of the wires on the steps. The z coordinates of the cursor were linearly fit to the known depths of the wires to obtain the calibration line. This calibration was performed for each of the image magnification/zoom conditions discussed below. While this calibration method is accurate and highly reproducible, it relies on the ability of the user to match the positions of the cursors and fibrils in the image and is therefore subjective. A future improvement of the calibration method would entail developing a computer program to determine the positions of the fibrils (e.g., their centers of masses).

F. True fibril depths

Through a careful examination of the fibril phantom, we noticed that minor warping of the sandwiched Lexan plates could cause the actual depths of the fibrils to differ from their nominal 1, 3, 5, 7, 9, and 11 mm values. To more accurately determine the true depths of the fibrils for each phantom layer configuration, we applied the calibration method described above to one set of the stereo pair images. The $\pm 6°$ magnification image pair was selected because the displacements in the fibril locations between the left- and right-eye images are the greatest for this image pair. The larger displacement results in greater localization accuracy since the limitation of approximately ± 1 pixel uncertainty in placement of the stereo cursors on the fibrils in the images will correspond to a smaller uncertainty in the actual depth. A third observer who was different from the two who participated in the observer study described below viewed, without the stereo glasses, the $\pm 6°$ magnification stereo pair images of the phantoms in the three multilayer configurations studied and adjusted the left- and right-eye cursor components to overlay each fibril. The measured z values were then converted to true depths in millimeters using the calibration lines derived with the step wedge phantom. In performing this procedure, we found that when the nominal 1 mm depth fibrils (i.e., those that were a distance of 1 mm from the bottom of the phantom) were viewed without the glasses, they were too close to each other for the accurate positioning of overlaying cursors. We therefore could not determine their true depths and only analyzed the true depths of the fibrils at the nominal 3, 5, 7, 9, and 11 mm depths in the phantoms. Hence, the results in this paper are only presented at those depths. This is a consequence of the method that was employed to determine the true depths and would not be a limi-

tation for a test object that was perfectly flat at each level since the true depths would then be equal to the nominal depths.

G. Observer study

For the observer experiment, we arranged the multilayered phantom in one configuration, and had two participants use the stereoscopic virtual cursor system to measure the depths of the fibrils in stereomammograms of the phantom for the seven stereo angle/geometry/display conditions described in the following: (1) ±3° stereo, contact geometry, no zoom; (2) ±3° stereo, contact geometry, zoom =2×; (3) ±6° stereo, contact geometry, no zoom; (4) ±6 stereo, contact geometry, zoom =2×; (5) ±3° stereo, geometric magnification =1.8×, no zoom; (6) ±3° stereo, geometric magnification =1.8×, zoom =2×; and (7) ±6° stereo, geometric magnification =1.8×, no zoom.

In addition, one of the observers repeated the entire study for two other phantom configurations. That observer also made a second set of measurements on one of the images (contact geometry with ±6° stereo shift angle) 8 months after the initial reading to assess reproducibility.

The stereo acuity of both observers was tested using a standard Randot® Circles Stereo test (Stereo Optical Co., Inc., Chicago, IL). In this test, the subject views a set of ten objects on the test pattern through polarized glasses. Each object consists of three circles, one of which when viewed stereoscopically should appear to be closer to the observer than the other two. The test subject is asked to identify the circle that appears closest in each object. Both observers in our study accurately identified each circle that was closer to them for all cases, indicating their level of stereopsis is at least 20 s of arc at a viewing distance of 16 in. Their performance is comparable to the average (21.3 s of arc) that has been measured with this test pattern for adults with excellent, balanced monocular visual acuity (at least 20/20 in each eye and equal acuity in both eyes).[16]

The cursor that was employed in our studies was a black cross-shaped cursor. It was symmetrically shaped with an overall height of 64 pixels, and an overall width of 64 pixels. The lines were two pixels thick, and the arrowheads at each end of the lines were three pixels long. In a brief preliminary study, it was found that a cursor like this that has vertical and horizontal lines worked best for measuring the depths of the +45° and −45° oriented fibrils.

The observers recorded their measured z coordinates of each fibril, and these z coordinates were converted into depths (or actually distances in front of the back surface of the phantom) using calibration lines that were calculated from the step wedge data discussed above. These depths were then compared with the known depths both by performing linear least-squares fits and by computing the rms and mean errors. Parameters of linear least-squares fits that were compared for each observer for the seven stereo angle/geometry/display zoom combinations included the slope, intercept, correlation coefficient (r value), and standard error of the estimate (SEE). The rms errors were computed using the equation

$$\text{rms error} = \sqrt{\frac{\Sigma_{i=1}^{N}(\text{true depth}_i - \text{measured depth}_i)^2}{N}},$$

when N was the total number of fibrils excluding those at the nominal 1 mm depth (N was equal to 40 for one of the phantom configurations and 42 for the other two phantom configurations). Finally, two-tailed paired t tests of the differences between the measured and true z coordinates of the fibrils for the various stereo imaging/viewing techniques were performed to determine the statistical significance of those differences.

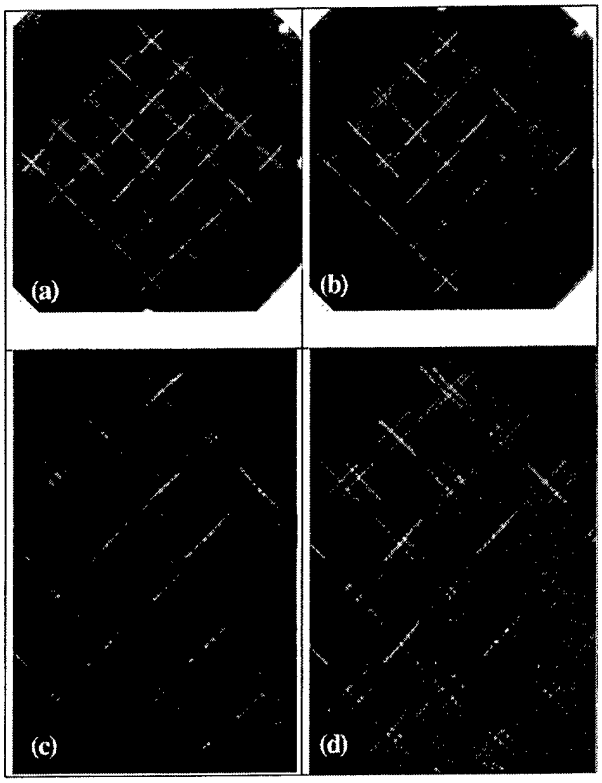


FIG. 2. Examples of images for (a) ±3° stereo shift angle, contact geometry, (b) ±6° stereo shift angle, contact geometry, (c) ±3° stereo shift angle magnification geometry, and (d) ±6° stereo shift angle magnification geometry. Note the combined images in this figure were created for illustration purposes by averaging the left- and right-eye images. In the actual stereo display, the left- and right-eye images are perceived individually. The averaging process helps illustrate the two images at once, but reduces the true image contrast.

III. RESULTS

Combined left-eye and right-eye images of the same phantom obtained with the following geometries: ±3° stereo shift—contact, ±6° stereo shift—contact, ±3° stereo shift—magnification, and ±6° stereo shift—magnification are shown in Fig. 2. Examples of the calibration lines that were computed from the step wedge measurements are shown in Fig. 3. The calibration lines for the ±3° and ±6° stereo shift angles in contact geometry are compared in part (a) of this figure, and the corresponding lines for the magnification ge-

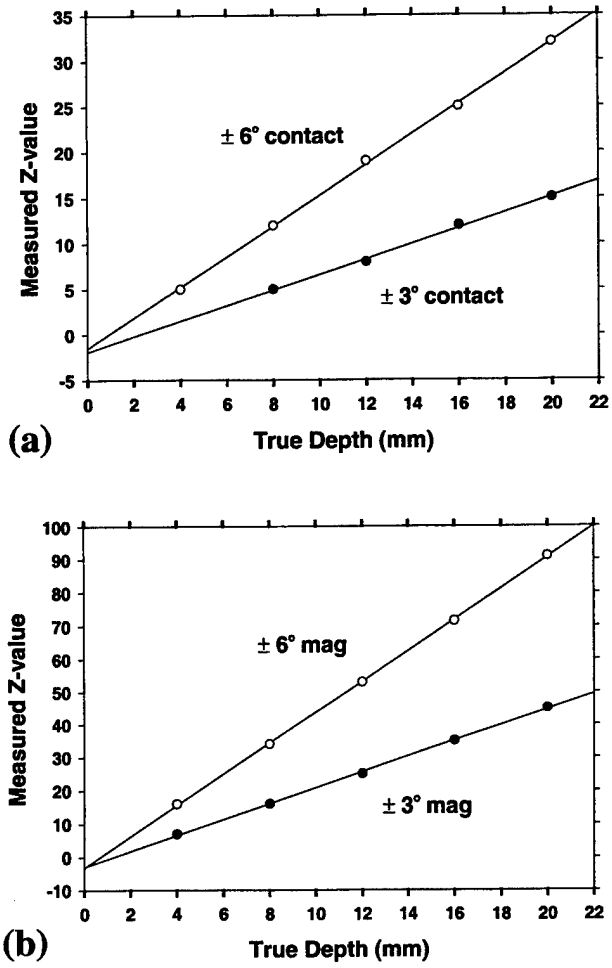


FIG. 3. (a) Calibration lines for ±3° stereo shift contact and ±6° stereo shift contact acquisition; (b) calibration lines for ±3° stereo shift magnification and ±6° stereo shift magnification acquisition.

ometry are compared in part (b). For the 2×-zoom mode, our cursor software effectively incremented the z values in steps of 0.5 per pixel shift instead of steps of 1 for the nonzoom mode. Consequently, the calibration lines that were computed in the 2×-zoom display modes were nearly identical to those in the nonzoom modes and they are therefore not shown in the figures. The calibration equations for the conversion of the measured z values to the depths (distances from the back of the phantoms) for the various imaging conditions are listed in Table I, below.

TABLE I. Calibration equations for converting measured z values to depths (distances in front of the backside of the phantom). *Note:* these calibration equations were derived by linearly fitting the data acquired from the step wedge images. Examples of the lines for the step wedge data are shown in Fig. 1.

Stereo Shift Angle	Geometry	Display zoom	Equation
±3 degrees	Contact	None	$\text{Depth (mm)} = \frac{Z+1.9}{0.85}$
±3 degrees	Contact	2×	$\text{Depth (mm)} = \frac{Z+1.55}{0.825}$
±6 degrees	Contact	None	$\text{Depth (mm)} = \frac{Z+1.5}{1.675}$
±6 degrees	Contact	2×	$\text{Depty (mm)} = \frac{Z+2.5}{1.75}$
±3 degrees	Magnification	None	$\text{Depth (mm)} = \frac{Z+2.9}{2.375}$
±3 degrees	Magnification	2×	$\text{Depth (mm)} = \frac{Z+2.6}{2.35}$
±6 degrees	Magnification	None	$\text{Depth (mm)} = \frac{Z+3.15}{4.688}$

Results comparing the performances of the two observers for the various stereo angle, geometric magnification, and display zoom combinations of the experiment are summarized in Tables II–IV. Examples of plots of the measured versus true depths for one observer for each of the imaging conditions are shown in Fig. 4. Errors in the depth measurements for each imaging condition for the one observer who viewed images in two additional phantom configurations are listed in Table IV. The paired t-test results for the observer who measured the z values of the fibrils in the phantom in three separate phantom layer configurations are listed in Table V. (Note, the paired t-test results for the other observer who made measurements of the fibrils in the first phantom configuration were very similar to those listed in part A of this table and are therefore not shown.)

The reproducibility of the observer's measurements in the same image read twice was excellent. For the 40 fibrils that were analyzed in the image, 35 of the measured z values were the same for both readings, and the remaining 5 z values differed by ±1. This translates to a rms difference in the z values of 0.354 (i.e., $\sqrt{5/40}$), which is equal to 0.21 mm.

TABLE II. Linear fit parameters for measured versus true depths of fibrils for two observers (values for observer 1 are indicated by subscript 1, and for observer 2 by subscript 2).

	3°	3° zoom	6°	6° zoom	3° mag	3° mag zoom	6° mag
r_1	0.956	0.949	0.994	0.996	0.982	0.978	0.998
r_2	0.934	0.921	0.994	0.996	0.985	0.981	0.998
intercept_1	0.669	0.030	−0.757	−0.306	−0.822	−0.470	−0.013
intercept_2	1.379	0.930	−0.910	−0.502	−0.777	−0.476	−0.082
slope_1	0.969	1.041	1.000	0.978	1.031	1.007	0.994
slope_2	0.961	1.002	1.037	0.997	1.049	1.016	1.004
SEE_1	0.884	1.036	0.331	0.263	0.604	0.656	0.170
SEE_3	1.103	1.264	0.329	0.254	0.560	0.606	0.174

TABLE III. Root mean square (rms), mean, and standard deviations of depth errors in mm (values for observer 1 are indicated by subscript 1, and for observer 2 by subscript 2).

	3°	3° zoom	6°	6° zoom	3° mag	3° mag zoom	6° mag
RMS_1	0.963	1.077	0.820	0.550	0.839	0.766	0.179
RMS_2	1.520	1.553	0.706	0.582	0.702	0.692	0.179
Mean $error_1$	0.420	0.354	−0.754	−0.482	−0.591	−0.421	−0.064
Mean $error_2$	1.068	0.945	−0.619	−0.527	−0.417	−0.357	−0.053
Std dev_1	0.877	1.030	0.326	0.268	0.603	0.648	0.169
Std dev_2	1.095	1.247	0.343	0.251	0.571	0.600	0.173

The computed rms errors (relative to the true depths) for the two independent readings were essentially identical (0.706 mm).

IV. DISCUSSION

The average rms errors for the two observers reading the same sets of images (Table III) were 1.2 mm (±3° contact, no zoom), 1.3 mm (±3° contact, zoom), 0.8 mm (±6° contact, no zoom), 0.6 mm (±6° contact, zoom), 0.8 mm (±3° magnification, no zoom), 0.7 mm (±3° magnification, zoom), and 0.2 mm (±6° magnification, no zoom). Corresponding values for the one observer who viewed all images of the phantom in three different layer configurations (Tables III and IV) were 1.6, 1.5, 0.7, 0.7, 0.5, 0.6, and 0.2 mm, respectively. In general, better results were obtained for the larger shift angle and for magnification geometry. Both of these improvements are in agreement with the analysis by Jiang *et al.* of the theoretical trends for stereo localization accuracy.[17] According to Jiang *et al.*, depth (z) localization error is inversely proportional to the x-ray tube shift (which is approximately equal to the tangent of the shift angle in our case), and directly proportional to the square of the focal spot-to-object distance (i.e., the error is smaller when the object is closer to the focal spot as in magnification geometry). Jiang *et al.* also showed that the sensitivity of the measurements to small changes in depth is proportional to the tube shift, and inversely proportional to the square of the focus-to-object distance.

Combining both observers' experimental results for all contact geometry images, we found that increasing the stereo shift angle from the conventional ±3° to ±6° improved the depth measurement accuracy by factors of about 1.2 to 4.0. The theoretical improvement considering only geometric factors[17] is approximately equal to the ratio of the tangents of the stereo shift angles which is 2.0 for this case (=tan 6°/tan 3°).

The variability in our results may be due to other sources of error. These include errors in the x-ray tube and slider positions for the acquisition of both the test phantom and calibration step wedge images, errors in the matching of the fiducial marker positions in the images for the desired 0 displacement at the bottom of the phantom, errors due to the limitation of a minimum 1 pixel (as opposed to fractional pixel) increment in the positioning of the virtual cursor for the z-value measurements in the test phantom and calibration phantom images, and uncertainties in the readers' determination of the cursor depth to overlay the fibril. The error in the measured depth due to the 1-pixel increment in the virtual cursor position is a function of the imaging geometry and can be computed by taking the derivative of the slope of the calibration line (Table I). This error is 1.18 mm for the ±3° stereo shift-contact geometry and 0.60 mm for the ±6° stereo shift-contact geometry. Thus, a change of 1 z-value unit in an observer's depth measurement of a fibril has a much greater effect for ±3° contact than for 6° contact geometry.

Comparing all measurements in the magnification versus contact geometries, we found improvements in depth accuracy by factors of 1.1–10.2 for ±3° magnification versus ±3° contact and by factors of 3.0–4.6 for ±6° magnification versus ±6° contact. The theoretical improvement due to geometrical factors in both cases is approximately proportional to the square of the ratio of the focus-to-object distances.[17] Considering a fibril located at about the midplane of the phantom (5 mm from the bottom of the phantom), and using the actual focus-to-magnification stand distance of 39.6 cm, and the focus-to-detector distance of 66 cm (assuming, for simplicity, that in the contact mode, the phantom is directly

TABLE IV. Root mean square (rms), mean, and standard deviations of depth errors in mm for two additional images read by observer 2.

	3°	3° zoom	6°	6° zoom	3° mag	3° mag zoom	6° mag
$RMS_{image\ 2}$	0.813	0.572	0.632	0.782	0.641	0.763	0.174
$RMS_{image\ 3}$	2.483	2.422	0.624	0.854	0.244	0.233	0.189
Mean $error_{image\ 2}$	0.691	0.303	0.373	0.620	−0.218	−0.271	0.031
Mean $error_{image\ 3}$	2.239	2.045	0.567	0.829	0.057	0.027	0.086
Standard $dev_{image\ 2}$	0.433	0.490	0.515	0.482	0.610	0.722	0.174
Standard $dev_{image\ 3}$	1.086	1.314	0.263	0.208	0.240	0.234	0.170

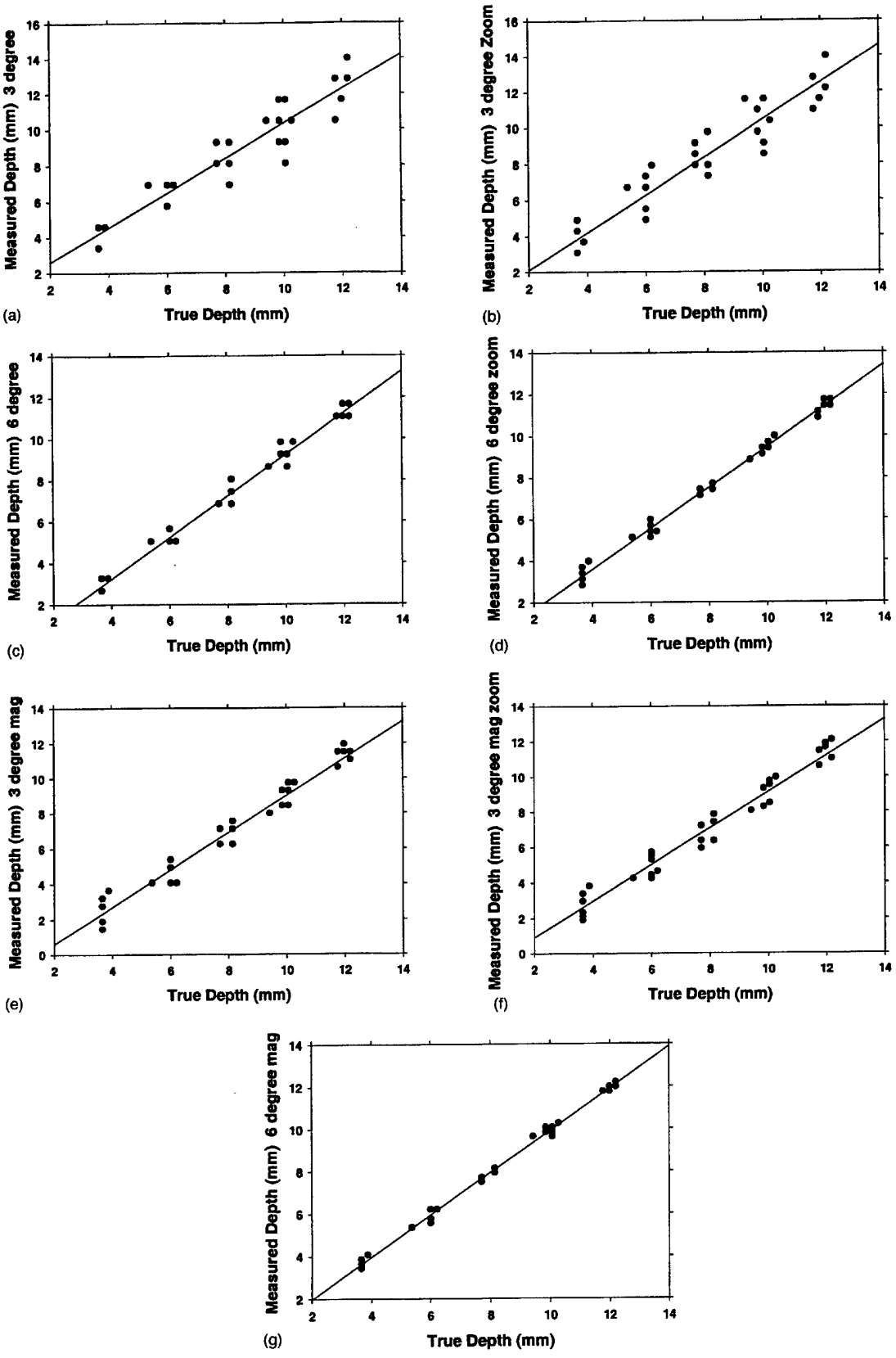


FIG. 4. Examples of measured versus true depths (distances from the back of the phantom) for observer number 1. (a) Stereo shift angle $=\pm 3°$, contact, no zoom; (b) stereo shift angle $=\pm 3°$, contact, zoom=2; (c) stereo shift angle $=\pm 6°$, contact, no zoom; (d) stereo shift angle $=\pm 6°$, contact, zoom=2; (e) stereo shift angle $=\pm 3°$, magnification mode, no zoom; (f) stereo shift angle $=\pm 3°$, magnification mode, zoom=2; (g) stereo shift angle $=\pm 6°$, magnification mode, no zoom.

TABLE V. p values for the paired t tests comparing the differences between the observed and measured z coordinates of the fibrils for the various stereoscopic imaging and display techniques (values are for three phantom configurations (different fibril patterns) analyzed by observer 2 (*=not significant since $p>0.05$; the rest are significant; *note:* the value of 0.000 00 means $p<10^{-5}$).

A Phantom layer configuration #1

	3° contact zoom	6° contact	6° contact zoom	3° mag	3° mag zoom	6° mag
3° contact	**0.22191*a**	0.00000	0.00000	0.00000	0.00000	0.00000
3° contact zoom		0.00000	0.00000	0.00002	0.00005	0.00001
6° contact			0.04878	0.03859a	0.01282	0.00000
6° contact zoom				**0.272 75*a**	**0.10784*a**	0.00000
3° mag					0.03496	0.00039
3° mag zoom						0.00416

a indicates not significant ($p>0.05$) for observer 1

B. Phantom layer configuration #2

	3° contact zoom	6° contact	6° contact zoom	3° mag	3° mag zoom	6° mag
3° contact	0.00000	0.00123	**0.462 20***	0.00000	0.00000	0.00000
3° contact zoom		**0.27320***	0.00008	0.00046	0.00066	0.00152
6° contact			0.00000	0.00031	0.00045	0.00012
6° contact zoom				0.00000	0.00000	0.00000
3° mag					**1.14830***	0.00541
3° mag zoom						0.00538

C. Phantom layer configuration #3

	3° contact zoom	6° contact	6° contact zoom	3° mag	3° mag zoom	6° mag
3° contact	0.01394	0.00000	0.00000	0.00000	0.00000	0.00000
3° contact zoom		0.00000	0.00000	0.00000	0.00000	0.00000
6° contact			0.00000	0.00000	0.00000	0.00000
6° contact zoom				0.00000	0.00000	0.00000
3° mag					**0.25839***	**0.45679***
3° mag zoom						**0.11153***

on top of the detector, and the fulcrum is in the plane of the detector), the theoretical improvement in depth accuracy is about a factor of 2.8. Variations in the observed improvements can be attributed to the reasons listed above.

In three of four cases, the measurement accuracies for ±3° magnification geometry were nearly identical to those for ±6° contact geometry. In one case, the accuracy was significantly superior (0.24 mm vs 0.62 mm) in ±3 magnification as compared to ±6° contact. The theoretical improvements due to geometrical considerations for this comparison is about 1.4. (As discussed above, the ±6° contact results should be better than the ±3° contact by a factor of 2 due to the increased tube shift, and the ±3° magnification results should be better than the ±3° contact by a factor of 2.8 due to the reduced focus-to-object distance. Combining these factors, the ±3° magnification measurements should be more accurate than ±6° contact by a factor of 1.4 (=2.8/2). The errors due to one-unit variations in the measured z values for the two geometries are similar (0.6 mm for ±6° contact and 0.42 mm for ±3° magnification.) Considering these factors and the other sources of measurement error discussed previously, the lack of observed improvement for the ±3° magnification geometry as compared to the ±6° contact geometry in most cases is not surprising. It should be noted that the rms errors for both geometries are very small (less than 1 mm) so either geometry would be adequate for making depth measurements.

Zooming the stereo images by a factor of 2 did not seem to improve the depth measurement accuracy. Although sometimes it appeared to improve the accuracy (7 of 12 times), other times it reduced the accuracy (5 of 12 times). There was basically no improvement on average (the difference in the average rms error for no zoom versus zoom, was 0.14 mm). The small differences one way or the other were there-

fore likely caused by experimental uncertainties due to the many factors described above. Thus, artificially increasing the displacement between objects viewed in the left- and right-eye stereo pairs through zooming the display does not have the same effect as increasing the displacement via increasing the stereo shift angle or increasing the geometric magnification. The stereo effect was more readily visualized in the zoom mode, but the signal-to-noise ratios of the fibril images were basically the same as those in the images displayed without zoom. In contrast, the acquisition of images with geometric magnification actually improves the signal-to-noise ratio.[18] The 2× zoom that was employed in this study was achieved by pixel replication. Results may be different for interpolative zoom. Finally, the use of greater zoom factors was also not explored. However, based on the results of a recently published study, increasing the display zoom factor may not be beneficial.[19] In that study, the observers' stereoacuities and depth perceptions were compared using a standard Randot stereotest pattern, with and without magnification via a 4× optical loupe and a 16× microscope. The researchers found that stereo acuity and depth perception decreased with increasing optical magnification of the pattern.

The paired *t*-test results listed in Table V indicate the majority of the differences between the depth measurement accuracies obtained with the ±3° and ±6° stereo shift angles, magnification and contact geometries, and normal and zoom displays are statistically significant ($p<0.05$). It is interesting to note that there was little consistency between the categories of the small number of insignificant differences ($p>0.05$) for the images created with the three different phantom configurations (parts A, B, and C of Table V). The only consistent insignificant result was that in two of the three cases, the accuracies for the ±3° mag and ±3° mag zoom depth measurements were not statistically significant. The *t*-test results for the two observers were quite similar for phantom configuration #1 (see footnote "a" of Table V, part A). Increasing the number of observers would have increased the statistical power of this study; however, the variability in the results due to the various factors described above would not have been reduced. Therefore the conclusions would likely be the same.

The best accuracy of about 0.2 mm, among all cases, was obtained with geometric magnification using a stereo angle of ±6°. Therefore, this is the mode we recommend for obtaining accurate depth measurements with virtual cursors in stereomammograms. Conventional stereoradiography is performed in contact mode using a stereo shift angle of ±3°. According to Christensen,[20] this angle was determined empirically "by trial and error." The angle is a compromise between the improved stereoscopic effect with increasing angle, the increased eye strain and difficulty in fusing the left- and right-eye images, especially at larger stereo angles, and the reduced patient coverage at larger stereo angles. For our particular case, the increased stereoscopic effect associated with the use of twice the conventional stereo angle and the use of ~1.8× geometric magnification did not cause undue eye strain. It did reduce the imaged field of view and this geometry could not be used for imaging an entire breast unless a larger-area detector was employed. Based on our measurements and observations, we recommend for stereoscopic imaging and depth measurements within an entire breast with the GE full-field digital mammography system, that a contact geometry be employed using a stereoshift angle of ±6° instead of the conventional ±3°. The overall results and recommendations of this study may differ for detectors having different pixel sizes and noise properties and for displays with different noise and contrast. A further investigation of the effects of these factors is warranted.

Finally, the depth measurement accuracies of the two observers in this study (Table III) were almost identical (in nearly all cases they were within 0.1 mm of each other). Both observers had excellent stereo acuity, and it would be expected that others with similar stereo acuity could achieve similar results after a period of training. Our previous studies with other observers have shown that there is a wide range of accuracies for depth measurements, especially for horizontally oriented objects.[3] All of the fibrils in this study had horizontal and vertical components, so it would be expected that even observers who have difficulties measuring the depths of horizontally oriented objects would be able to measure the depths of the diagonally oriented fibrils, although their accuracies may not be as high. Since it is unlikely that fibrous tissues or spicules from masses in mammograms are exactly horizontal to the stereo shift direction, the angulation would enable reasonably accurate measurements in clinical images. However, the inhomogeneous anatomical background within the clinical images may partially obscure the fibrils, which would increase the difficulty of making accurate depth measurements with the stereo cursor. It is possible that with additional training and the use of depth cues (e.g., 3-D wire boxes placed about the objects of interest, or 3-D rulers in the image) in clinical images, the depth measurement accuracies of most viewers would be adequate. The development of such depth cues will be the subject of our future investigations.

ACKNOWLEDGMENTS

This work is supported by U.S. Army Medical Research and Material Command Grants No. DAMD 17-98-1-8210 and No. DAMD 17-99-1-9294. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred.

[a]Corresponding author: Department of Radiology, University of Michigan Hospitals, Room B1 F510C, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-0030. Office 734-936-7474; Fax: 734-936-7948; Electronic mail: goodsitt@umich.edu

[1] D. M. Chelberg, J. Hsu, C. G. Babbs, Z. Pizlo, and E. J. Delp, "Digital stereomammography," in *Proceedings of the 2nd International Workshop on Digital Mammography*, Excepta Medica International Congress Series 1069 (Elsevier Science, York, England, 1994), pp. 181–190.

[2] J. Hsu, D. M. Chelberg, C. F. Babbs, Z. Pizlo, and E. J. Delp, "Preclinical ROC studies of digital stereomammography," IEEE Trans. Med. Imaging **14**, 318–327 (1995).

[3] M. M. Goodsitt, H.P. Chan, and L. M. Hadjiiski, "Stereomammography: Evaluation of depth perception using a virtual 3D cursor," Med. Phys. **27**, 1305–1310 (2000).
[4] D. J. Getty, R. M. Pickett, and D. J. D'Orsi, "Stereoscopic digital mammography: improving detection and diagnosis of breast cancer," in *Computer Assisted Radiology and Surgery 2001*, Proceedings of the 15th International Congress and Exhibition, Berlin, 27–30 June 2001, edited by H. U. Lemke, M. W. Vannier, K. Inamura, A. G. Farman, and D. Doi (Elsevier, Amsterdam, 2001; International Congress Series 1230), pp. 506–511.
[5] M. M. Goodsitt, H. P. Chan, J. M. Sullivan, K. L. Darner, and L. M. Hadjiiski, "Evaluation of the effect of virtual cursor shape on depth measurements in digital stereomammograms," in *IWDM 2000 5th International Workshop on Digital Mammography*, edited by M. J. Yaffe (Medical Physica Publishing, Madison, 2001), pp. 45–50.
[6] H. P. Chan, M. M. Goodsitt, K. L. Darner, J. M. Sullivan, L. M. Hadjiiski, N. Petrick, and B. Sahiner, "Effects of stereoscopic imaging technique on depth discrimination," in Ref. 5, p. 13–18.
[7] B. Leduc, R. Shumak, G. E. Mawdsley, M. J. Yaffe, and R. A. Jong, "Merits of using a 3-D visualization technique to target microcalcifications," in Ref. 5, pp. 56–64.
[8] L. T. Niklason *et al.*, "Digital tomosynthesis in breast imaging," Radiology **205**, 399–406 (1997).
[9] R. L. Webber, H R. Underhill, and R. I. Freimanis, "A controlled evaluation of tuned-aperture computed tomography applied to digital spot mammography," J. Digit. Imaging **13**, 90–97 (2000).
[10] S. Suryanarayanan, A. Karellas, S. Vedantham, S. P. Baker, S. J. Glick, C. J. D'Orsi, and R. L. Webber, "Evaluation of linear and nonlinear tomosynthetic reconstruction methods in digital mammography," Acad. Radiol. **8**, 291–224 (2001).
[11] V. Raptopoulos, J. K. Baum, M. Hochman, A. Karellas, M. J. Houlihan, and C. J. D'Orsi, "High resolution CT mammography of surgical biopsy specimens," J. Comput. Assist. Tomogr. **20**, 179–184 (1996).
[12] J. M. Boone, T. R. Nelson, K. K. Lindfors, and J. A. Seibert, "Dedicated breast CT: Radiation dose and image quality evaluation," Radiology **221**, 657–667 (2001).
[13] S. L. Warren, "Roentgenologic study of the breast," Am. J. Roentgenol. **24**, 113–124 (1930).
[14] A. D. A. Maidment, P. Bakic, and M. Alberg, "Is stereomammography possible without increasing dose?," *The 6th International Workshop on Digital Mammography*, 22–25 June 2002, Bremen, Germany, Program & Abstracts, 2002, p. 27.
[15] S. Vedantham, A. Karellas, S. Suryanarayanan, D. Albagli, S. Han, E. J. Tkaczyk, C. E. Landberg, B. Opsahl-Ong, P. R. Granfors, I. Levis, C. J. D'Orsi, and R. E. Hendrick, "Full breast digital mammography with an amorphous silicon-based flat-panel detector: physical characteristics of a clinical prototype," Med. Phys. **27**, 558–567 (2000).
[16] K. Simons, "Stereoacuity norms in young children," Arch. Opthalmol. **99**, 439–445 (1981).
[17] H. Jiang, H. Liu, G. Wang, W. Chen, and L. L. Fajardo, "A localization algorithm and error analysis for stereo x-ray image guidance," Med. Phys. **27**, 885–893 (2000).
[18] K. Doi and H. Imhof, "Noise reduction by radiographic magnification," Radiology **122**, 479–487 (1977).
[19] L. T. Du, L. F. Wessels, J. P. Underdahl, and J. D. Auran, "Stereoacuity and depth perception decrease with increased instrument magnification: comparing non-magnified system with lens loupes and a surgical microscope," Binocul. Vis. Strabismus. Q. **16**, 61–67 (2001).
[20] T. S. Curry, J. E. Dowdey, and R. C. Murry, *Christensen's Physics of Diagnostic Radiology*, 4th ed. (Lea & Febiger, Philadelphia, 1990).

CE: BM

INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. 48 (2003) 1–14 PII: S0031-9155(03)63558-X

Effects of magnification and zooming on depth perception in digital stereomammography: an observer performance study

Heang-Ping Chan, Mitchell M Goodsitt, Lubomir M Hadjiiski, Janet E Bailey, Katherine Klein, Katie L Darner and Berkman Sahiner

Department of Radiology, University of Michigan, Ann Arbor, MI 48109, USA

E-mail: chanhp@umich.edu

Received 15 May 2003
Published DD MMM 2003
Online at stacks.iop.org/PMB/48/1

Abstract

We are evaluating the application of stereoscopic imaging to digital mammography. In the current study, we investigated the effects of magnification and zooming on depth perception. A modular phantom was designed which contained six layers of 1 mm thick Lexan plates, each spaced 1 mm apart. Eight to nine small, thin nylon fibrils were pasted on each plate in horizontal or vertical orientations such that they formed 25 crossing fibril pairs in a projected image. The depth separation between each fibril pair ranged from 2 to 10 mm. A change in the order of the Lexan plates changed the depth separation of the two fibrils in a pair. Stereoscopic image pairs of the phantom were acquired with a GE full-field digital mammography system. Three different phantom configurations were imaged. All images were obtained using a Rh target/Rh filter spectrum at 30 kVp tube potential and a ±3° stereo shift angle. Images were acquired in both contact and 1.8X magnification geometry and an exposure range of 4 to 63 mAs was employed. The images were displayed on a Barco monitor driven by a Metheus stereo graphics board and viewed with LCD stereo glasses. Five observers participated in the study. Each observer visually judged whether the vertical fibril was in front of or behind the horizontal fibril in each fibril pair. It was found that the accuracy of depth discrimination increased with increasing fibril depth separation and x-ray exposure. The accuracy was not improved by electronic display zooming of the contact stereo images by 2X. Under conditions of high noise (low mAs) and small depth separation between the fibrils, the observers' depth discrimination ability was significantly better in stereo images acquired with geometric magnification than in images acquired with a contact technique and displayed with or without zooming. Under our experimental conditions, a 2 mm depth discrimination was achieved with over 60% accuracy on contact images with and without zooming, and with over

0031-9155/03/000001+14$30.00 © 2003 IOP Publishing Ltd Printed in the UK

90% accuracy on magnification images. This study indicates that stereoscopic imaging, especially with magnification, may be useful for visualizing the spatial distribution of microcalcifications in a cluster and for differentiating overlapping tissues from masses on mammograms.

1. Introduction

At present, x-ray mammography is the only diagnostic procedure with a proven capability for detecting early stage, clinically occult breast cancers (Seidman *et al* 1987). Although mammography has a high sensitivity for detection of breast cancers when compared to other diagnostic modalities, studies indicate that radiologists do not detect all carcinomas that are visible on retrospective analyses of the images (Wallis *et al* 1991, Bird *et al* 1992, Harvey *et al* 1993, Beam *et al* 1996). These missed detections are often a result of the very subtle nature of the radiographic findings. However, one of the major deficiencies of mammography is its inability to discern lesions hidden behind dense fibroglandular tissues (Jackson *et al* 1993). It is estimated that about 20% of the breast cancers in dense breasts are not detected by mammography (Wallis *et al* 1991, Bird *et al* 1992). It is therefore important to improve the sensitivity of mammography in imaging dense breasts. With the advent of high resolution digital detectors, new breast imaging techniques such as stereomammography (Goodsitt *et al* 2000, 2002, Chan *et al* 2001, Getty *et al* 2001, Maidment *et al* 2003), digital tomosynthesis (Niklason *et al* 1997, Suryanarayanan *et al* 2000) and computed tomography (Raptopoulos *et al* 1996, Boone *et al* 2001) are being developed to alleviate this problem.

A conventional radiograph is a projection image. The anatomical structures along the x-ray beam path are projected onto a two-dimensional image plane and overlap with each other. The overlying tissue structures often obscure the visibility of subtle lesions of interest in a radiograph. The camouflaging effect of the anatomical structures is the main cause of missed diagnosis. Stereoscopic imaging will allow the overlying structures to be perceived at different depths, thereby reducing the camouflaging effect. It has been reported that digital stereomammography allowed the detection of additional lesions that were obscured on screen-film mammograms (Getty *et al* 2001).

Stereoscopic radiography has been attempted for different types of examinations (Doi *et al* 1981, Kelsey *et al* 1982, Doi and Duda 1983, Higashida *et al* 1988, Trocme *et al* 1990, Ragnarsson and Karrholm 1992). The principle of stereoscopic imaging is shown in figure 1. The x-ray focal spot is shifted, along a direction parallel to the image plane, to the left and the right of the central axis to obtain two images of the object. The object has to remain stationary during the process. The images are referred to as the left-eye (LE) and the right-eye (RE) images. When the two images are positioned properly and viewed by trained eyes or with the aid of a stereoscope so that the left eye sees only the LE image and the right eye sees only the RE image, the parallax between the two images creates the depth perception. It is important to match the amount of stereoscopic shift to the imaging geometry to obtain best depth perception with minimal eyestrain. In general, a larger stereoscopic shift produces improved depth perception; however, as the shift increases, it becomes more difficult for the observer to fuse the images for the stereoscopic effect, and eye fatigue increases. According to *Christensen's Physics of Diagnostic Radiology* (Curry *et al* 1992), early radiologists determined empirically that a tube shift of 10% of the source-to-film distance worked well. This translates to a stereo shift angle of about $\pm 3°$ ($\tan^{-1}(0.05) \cong 3°$) relative to a normal to the detector.

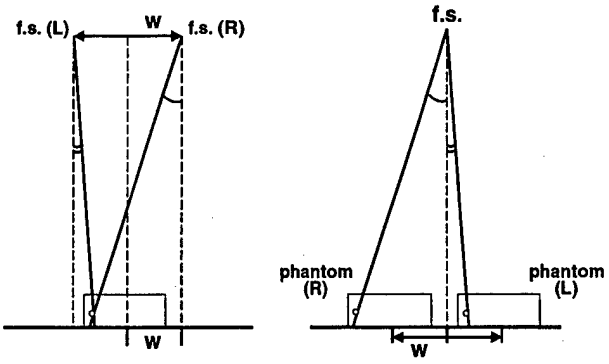


Figure 1. Imaging geometry for acquisition of stereoscopic image pairs. On the left is a conventional 'focal-spot-shift' method in which the focal spot is shifted to the left and to the right by a distance of w to expose the left-eye and right-eye images. On the right is an equivalent 'phantom-shift' method in which the phantom is shifted to the left and the right of the central ray by a distance of w. The 'phantom (R)' geometry corresponds to the 'f.s. (R)' geometry, and the 'phantom (L)' geometry corresponds to the 'f.s. (L)' geometry.

Stereoscopic imaging has not achieved widespread acceptance in clinical practice, mainly because of the doubled film cost and increased patient exposure (Curry *et al* 1992). A secondary problem is the need to train the eyes to perceive the stereoscopic effect without aid, or to use a special stereoscope, with careful arrangement of the films. Digital imaging may make the stereoscopic technique a viable approach because no additional film cost will be required. Furthermore, a digital detector has a wider linear-response range and a higher contrast sensitivity than a screen-film system so that good-quality images may be acquired at a reduced radiation dose. Images in digital form can be subjected to image processing, further enhancing the visibility of the image details. Digital stereoscopic images can be viewed on a single electronic display device that displays the left-eye and right-eye images alternately at a high refresh rate. The observer views the displayed images through a stereo goggle such as a pair of liquid crystal display (LCD) glasses. The LCD panels act as electronic shutters, blocking the light to the left eye and the right eye alternately. When the goggle is synchronized with the display so that only the left eye can see the left-eye image and only the right eye can see the right-eye image in rapid succession, the parallax between the two images will create a stereo effect. The stereoscopic images can be viewed singly in the conventional manner or stereoscopically on the same viewing station to provide complementary diagnostic information.

An additional advantage of stereoscopic imaging is the three-dimensional (3D) information it provides on the lesions of interest. It has been reported that the 3D distribution of microcalcifications may be correlated with the malignant or benign nature of the cluster (Conant *et al* 1996, Maidment *et al* 1996). Spiculations from a mass may be more readily distinguished from overlapping tissues under stereoscopic viewing conditions. This supplementary diagnostic information may improve the classification of malignant and benign lesions, thereby reducing unnecessary biopsies and increasing the positive predictive value of mammography.

We are evaluating the application of stereoscopic techniques to digital mammography. Previously we studied the effects of stereo shift and imaging conditions on the depth perception of fibrils in stereo phantom images (Chan *et al* 2001). We also demonstrated that a virtual cursor could provide accurate depth measurements in stereo phantom images acquired under

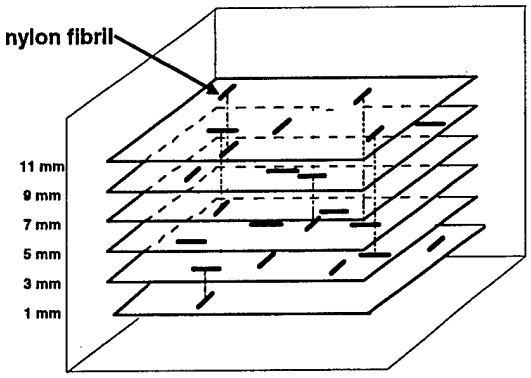


Figure 2. A modular stereo phantom consists of six layers of Lexan plates on which nylon fibrils were pasted either in the horizontal or the vertical direction in one of the 5 × 5 matrix locations on one of the plates. By properly arranging 50 fibrils on the six plates, the projected image will contain 25 pairs of crossing fibrils with different depth separations between the two fibrils.

mammographic conditions (Goodsitt *et al* 2000, 2002). In the present study, we further evaluated the effects of geometric magnification, display zooming and x-ray exposure on the visual depth discrimination of fibril-like objects in stereo phantom mammograms.

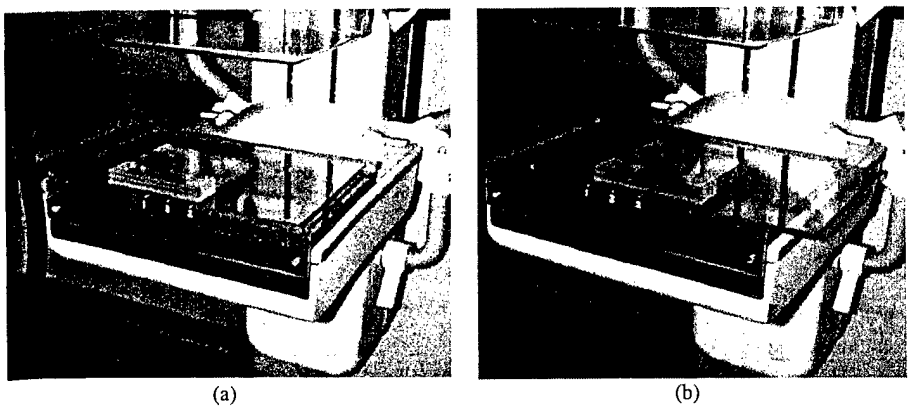
2. Materials and methods

2.1. Modular stereo phantom

We have designed a modular stereo phantom for evaluation of depth perception in stereomammography. A schematic of the phantom is shown in figure 2. The phantom consists of six 1 mm thick Lexan sheets, each separated by 1 mm thick spacers. Each Lexan plate contains a 5 × 5 array of object areas. Fifty nylon fibrils, each about 8 mm in length and 0.53 mm in diameter, are arranged in these object areas. Twenty-five fibrils are oriented perpendicular and another 25 are oriented parallel to the stereo shift direction. Henceforth, we will refer to the fibrils oriented in the perpendicular direction as 'vertical' and the fibrils oriented in the parallel direction as 'horizontal'. On average, eight to nine fibrils are placed in the object areas on each Lexan plate. The location and orientation of a fibril on a Lexan plate are randomly chosen with the constraint that, in the projection image, each object area contains the projection of one vertical and one horizontal fibril that cross each other. These plates can be arranged in different orders to produce many independent object configurations, i.e., the depth separation and whether the vertical fibril is in front of or behind the horizontal fibril in a given pair change when the order of the plates changes. Different types of test objects such as microcalcifications may be used in place of the fibrils to generate a different phantom although only fibrils were used in the present study. An additional 1 mm thick Lexan plate without objects is placed on top of the phantom to protect the test objects on the top layer. The total thickness of Lexan in the phantom is therefore 7 mm.

2.2. Image acquisition and display

Digital stereoscopic image pairs were acquired with a GE Senographe 2000D digital mammography system. The system employs a digital detector consisting of a CsI:Tl scintillator and an amorphous-Si active matrix flat panel. The detector measures 23 cm × 18 cm, with



(a) (b)

Figure 3. The phantom platform placed on top of the detector of the digital mammography system in the setup for contact imaging. The sliding plate with the stereo phantom was shifted to the left in (a) and to the right in (b) at a distance equivalent to a $\pm 3^\circ$ stereo angle.

a pixel size of 100 μm $\times$ 100 μm. We acquired stereoscopic pairs of images using x-ray techniques of 30 kVp, Rh target/Rh filter, a $\pm 3^\circ$ stereo angle, contact (reciprocating grid, 0.3 mm focal spot) and 1.8X magnification (no grid, 0.15 mm focal spot) geometries and exposures of 4, 8, 32 and 63 mAs per image. A relatively hard x-ray beam was used to produce lower contrast images, thus making the perception task more challenging.

Since the digital mammography system was not designed for stereo imaging, it is not easy to move the x-ray focal spot to the angulated positions at $\pm 3^\circ$ from the central ray for imaging the stereo pair. We used an equivalent imaging geometry (figure 1) in which the phantom was shifted instead of the x-ray source. Comparing the two geometries in figure 1, it can be seen that the position of the x-ray source relative to the phantom is the same except that the focal-spot-to-detector distance is slightly shorter in the phantom-shift geometry because the x-ray source moves along an arc. However, with the 3° shift about a fulcrum of rotation at a distance of 46 cm from the x-ray focal spot, this error is less than 0.1%. Using the geometry of the x-ray system, we calculated that a $\pm 3^\circ$ stereo shift of the x-ray focal spot is equivalent to a phantom shift distance of ± 2.4 cm from its central position. The shift distance is the same for either the contact geometry or the magnification geometry.

We built a phantom platform with Lexan (figure 3) to move the phantom from one position to the other in a direction parallel to the chest wall for acquisition of the stereo pair. The platform has a stationary base that is fitted on the Bucky holder for contact geometry or on the magnification stand for magnification geometry. A sliding plate on top of the base is used to support and move the phantom parallel to the chest wall direction of the mammography system. The sliding plate is guided by two guardrails parallel to the chest wall to ensure a precise and reproducible movement. The total thickness of the Lexan platform is 1.7 cm. During image acquisition, care was taken such that the phantom remained stationary on the sliding plate while the plate was moved between the two shifted positions. Pictures of the phantom platform at the left and right shifted positions in the contact geometry are shown in figure 3.

Three different configurations of the fibril phantoms were imaged under each exposure condition. Twenty-four (= 3 phantoms $\times$ 4 exposures $\times$ 2 geometries) stereo image pairs were thus produced. For each exposure and geometry condition, there were a total of 75 pairs of fibril images (25 fibril pairs in each phantom configuration $\times$ 3 configurations) at

BEST AVAILABLE COPY

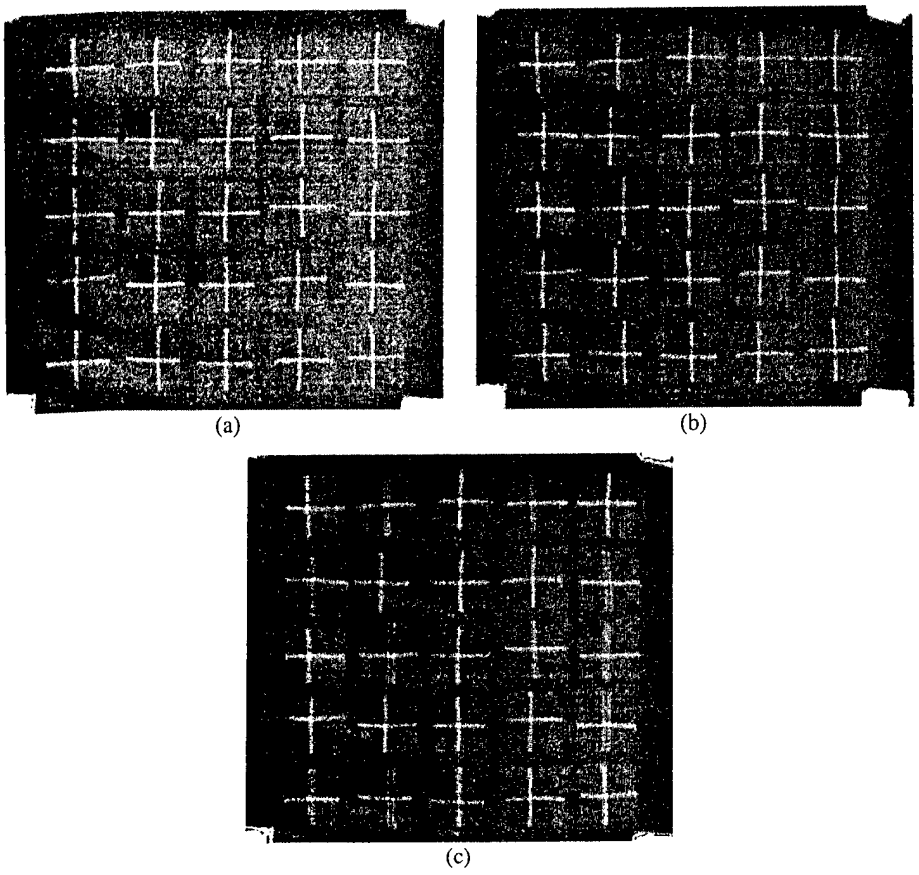


Figure 4. A stereo image pair acquired in contact geometry: (a) left-eye image, (b) right-eye image, (c) displayed stereo image—the 25 pairs of crossing fibrils showed different degrees of stereo shift due to the different depth separations between the two fibrils.

five different depth separations (2, 4, 6, 8, 10 mm) to be evaluated. An example of a stereo image pair acquired with contact geometry is shown in figure 4.

The images were displayed on a 21″ Barco-Metheus (Beaverton, OR) model 521 display monitor driven by a Barco-Metheus model 1760S stereoscopic board and a SUN Microsystems (Palo Alto, CA) Ultra 10 computer using in-house developed software. The Metheus board displays 1408 × 1408 × 8 bit images at a refresh rate of 114 Hz. It operates in a page flipping stereoscopic mode with the left- and right-eye images displayed alternately. NuVision (Beaverton, OR) LCD stereoscopic glasses were used for viewing the stereoscopic images. The stereo workstation is shown in figure 5.

2.3. Observer experiment

Five observers including two experienced mammographic radiologists participated in the experiment. All observers took a standard Randot Circles Stereo test (Stereo Optical Co., Inc., Chicago, IL) to evaluate their stereo acuity. In this test, the observers viewed ten rows of test

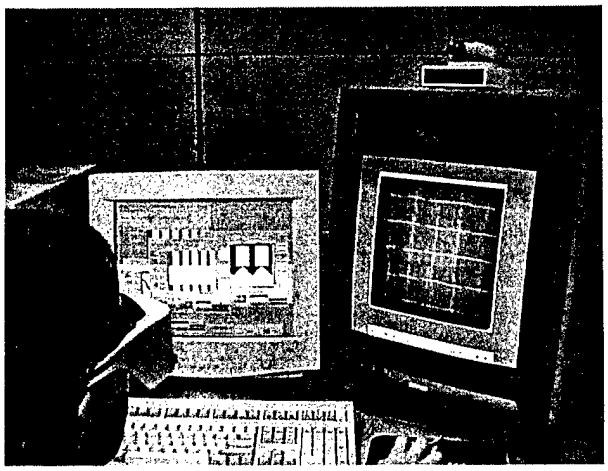


Figure 5. The SUN-based Barco-Metheus stereo display workstation used in the observer study.

objects through polarized glasses. Each row contains three objects, one of which is the object that exhibits the stereo effect and appears to be on a plane different from the other two objects. The difference progressively decreases due to decreasing stereo shift, and thus increasing the degree of difficulty. The observers had to determine which one was the stereo object in each row. All of our observers could correctly identify all the objects in the test, indicating that their stereo acuity was at least 20 s of arc at a viewing distance of 16″. Their performance is comparable to the average performance (21.3 s of arc) for adults with excellent, balanced monocular visual acuity measured with this test pattern (Simons 1981).

For our experiment, the task for the observers was to visually judge whether the vertical fibril in each pair of overlapping fibrils was in front of (closer to the observer) or behind the horizontal fibril. The observer was not informed of the truth after reading each image. Each observer read the images sequentially in a different randomized order. The reading order of the fibril pairs in the 5 × 5 matrix in a given image was changed randomly by changing the starting location and the reading direction, e.g., from the upper left corner and by row, from the upper left corner and by column, from the lower right corner and by row, etc in an effort to further reduce any potential effects of memorization. The contact images were read in two modes—in a regular and in a 2X-zoom mode, referred to as the contact and contact-zoom mode, respectively. The 1.8X geometric magnification images were read only in the regular mode, referred to as the magnification mode. The contact and the contact-zoom modes of the same image were not read consecutively so that it was unlikely that the observer would remember the results of the other mode. The observers were not informed of the conditions under which the image being viewed was acquired, although the difference in image sizes between the contact and the magnification or contact-zoom images was apparent. Prior to reading the test cases, the observers participated in training sessions to become familiar with the reading task.

3. Results

In this study, we quantified the accuracy of depth perception in stereo images by an observer as the percentage of correct decisions to differentiate the relative depths of the fibrils. The

BEST AVAILABLE COPY

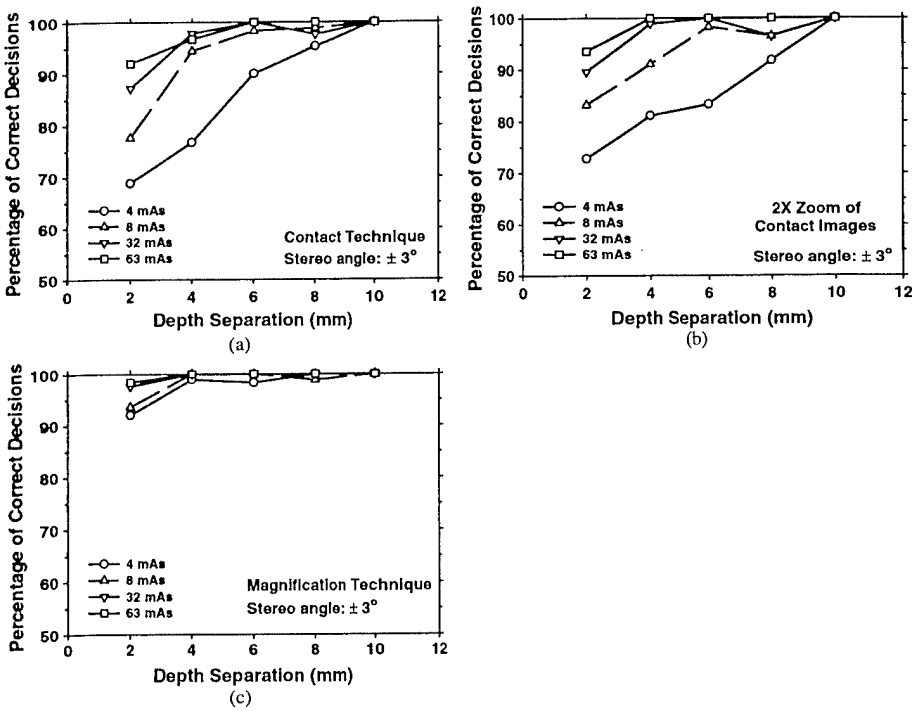


Figure 6. Percentage of correct decisions in differentiating the relative depths of two crossing fibrils, averaged over five observers. (a) Stereo images acquired with contact geometry, (b) stereo images acquired with contact geometry, displayed with 2X zoom and (c) stereo images acquired with 1.8X geometric magnification.

results were averaged over the five observers. The dependences of the average percentage of correct decisions on the depth separation of the crossing fibrils and the exposure levels for the contact, contact-zoom and magnification modes are shown in figures 6(a)–(c), respectively. Generally, the percentage of correct decisions increases as the exposure increases and as the depth separation between the two fibrils increases. The dependence on the depth separation diminishes as the exposure increases. At small depth separations, the noise in the images has a much stronger influence on the depth perception. The percentage of correct decisions increases rapidly as the exposure increases. At a depth separation of 10 mm, the percentage of correct decisions was 100% for all observers in all three modes, regardless of exposure levels. The curves for contact mode are very similar to those for the contact-zoom mode. On the other hand, the curves for the magnification mode are much higher than those for the contact or contact-zoom mode. The standard deviation of the percentage of correct decisions varied from 0 to about 13% for the contact mode and the contact-zoom mode, and varied from 0 to 6.7% for the magnification mode. The standard deviation generally increased as the percentage of correct decisions decreased. The standard deviation was zero for the conditions in which the average percentage of correct decisions was 100% because all observers were 100% correct in their decisions.

The statistical significance of the differences between every two modes, contact versus contact-zoom, contact versus magnification and contact-zoom versus magnification, was

Table 1. The p-values from the two-tailed paired t-test of the differences in the percentage of correct decisions in differentiating the depths of two crossing fibrils between the contact and the magnification modes. The entries indicated by '–' have p-values > 0.05 so that the differences are not statistically significant. The two entries indicated by –* have p-values of 0.089 and 0.065, close to being significant.

Depth separation (mm)	Exposure			
	4 mAs	8 mAs	32 mAs	63 mAs
2	0.007	0.000	–*	–
4	0.009	–*	–	–
6	0.034	–	–	–
8	–	–	–	–
10	–	–	–	–

Table 2. The p-values from the two-tailed paired t-test of the differences in the percentage of correct decisions in differentiating the depths of two crossing fibrils between the contact-zoom and the magnification modes. The entries indicated by '–' have p-values >0.05 so that the differences are not statistically significant.

Depth separation (mm)	Exposure			
	4 mAs	8 mAs	32 mAs	63 mAs
2	0.014	0.003	0.011	–
4	0.016	0.003	–	–
6	0.021	–	–	–
8	–	–	–	–
10	–	–	–	–

estimated by a two-tailed paired t-test. The paired t-test was performed for each depth separation and exposure level over the five observers. The p-values for the contact-versus-magnification and contact-zoom-versus-magnification for the various exposure and depth separation conditions are shown in tables 1 and 2, respectively. The differences between the contact and the magnification modes are statistically significant (p-value < 0.05) for depth separations of 2 mm to 6 mm at 4 mAs, and for a 2 mm separation at 8 mAs. The differences between the contact-zoom and the magnification modes are statistically significant for depth separations of 2 mm to 6 mm at 4 mAs, for 2 mm to 4 mm separations at 8 mAs and for a 2 mm separation at 32 mAs. The differences between the contact and contact-zoom modes are not statistically significant.

We further analysed separately the reading results for two subgroups of crossing fibrils, one subgroup with the vertical fibril in front of (referred to as the 'front group') and the other behind (referred to as the 'back group') the horizontal fibril. The average percentage of correct decisions as a function of depth separation was obtained for each exposure level, similar to the analysis for the entire group described above. We observed that the average percentage of correct decisions for the front group was often greater than that for the back group for a given exposure and a given depth separation. A histogram of the differences between the average percentage of correct decisions for the front group and that for the back group under the corresponding conditions is plotted in figure 7 for the 60 conditions studied (4 exposure levels × 5 depth separations × 3 modes). Of the 60 differences, only three were negative, i.e., the average percentage of correct decisions for the front group was smaller than that for the back group. All 28 zeroes (no difference) happened when both the front group and the back group were 100% correct. These occurred when the depth separations were large and the exposures

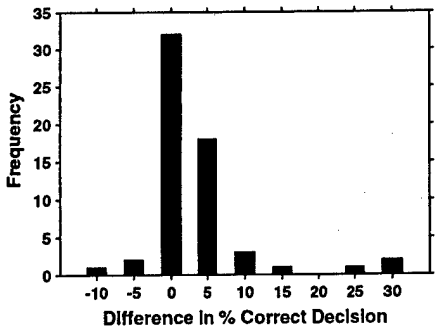


Figure 7. Histogram of the difference in the percentage of correct decisions between the front group (vertical fibril in front of the horizontal fibril) and the back group (vertical fibril behind the horizontal fibril).

were high. The other 29 differences were positive, i.e., the percentage of correct decisions for the front group was larger than that for the back group. However, paired *t*-tests of the differences between the front and the back groups under each condition did not reach statistical significance, probably due to the smaller numbers of samples when they were separated into subgroups.

4. Discussion

Depth perception in stereomammography depends on a number of factors including the x-ray exposure, depth separation and imaging geometry. In this study, we investigated the dependence of the percentage of correct decisions in differentiating the relative depths of two crossing fibrils on these factors using images acquired with $\pm 3°$ stereo angle. We found that a 2 mm depth resolution could be achieved with over 60% accuracy for all imaging conditions studied. For contact geometry, the accuracy improved to greater than 90% at higher exposures. Magnification stereomammography provided over 90% accuracy at a 2 mm depth resolution for all exposure levels studied. An interesting finding was that displaying the stereo images acquired with contact geometry in 2X-zoom mode did not improve the depth discrimination accuracy. In contrast, geometric magnification with about the same factor (1.8X) of enlargement significantly improved depth perception. This result indicated that the signal-to-noise ratio (SNR) might be a more important factor affecting depth perception than the perceived object size. When the images are acquired with geometric magnification, both the spatial resolution of the detector relative to the object and the x-ray quanta per unit object area increase (Doi and Rossmann 1974, Doi and Imhof 1977), although geometric unsharpness may somewhat reduce the gain in resolution. The improved detector resolution and reduced quantum mottle contributed to an increase in the SNR of the image. On the other hand, if the images were acquired in contact geometry and the object size was enlarged electronically by zooming during display, the inherent SNR of the object in the images remained fixed. The differences in the perceived SNRs (Loo *et al* 1985, Aufrichtig 1999) between the contact mode and the contact-zoom modes would be mainly caused by the change in the perceived noise of the display monitor relative to the object size as well as the change in the perceived signal and image noise spectra due to zooming relative to the frequency response of the observer's visual system. Display zooming may have a strong influence if the perception is limited by

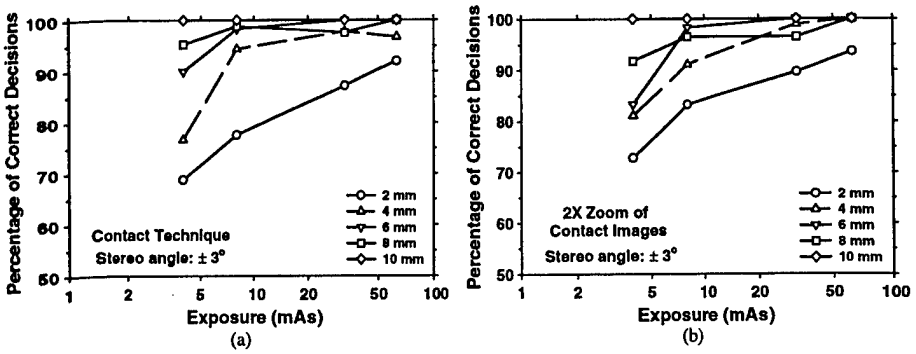


Figure 8. Dependence of the percentage of correction decisions on the logarithm of exposure. (a) Contact mode. (b) Contact-zoom mode.

the resolution of the human visual system. In our case, the objects were relatively large and it was unlikely that they were near the resolution limit of the observers' visual system. The improved SNR inherent in the geometrically magnified images may therefore have a dominant effect on depth discrimination compared with that of increased displayed object size. This may be the reason that display zoom did not have a strong impact on the accuracy of depth perception in this study.

We plotted the average percentage of correct decisions as a function of log exposure (mAs) for each depth separation between the two crossing fibrils in the pair as shown in figures 8(a) and (b), respectively, for the contact and contact-zoom modes. The average percentages of correct decisions for the magnification mode under all conditions were greater than 90% and were not plotted. Three of the data points at a 8 mm depth separation appear to be lower than those at a 6 mm depth separation for a given exposure. Since the standard deviation at each of these data points is about twice as large as the difference between the percentage of correct decisions at 6 mm and 8 mm separations, the apparent local reverse of the trend can likely be attributed to experimental uncertainties. From the curves of the contact and contact-zoom modes, it appears that there is no simple relationship between the percentage of correct decisions and log exposure. The curve shape varied among the different depth separations and they were not linear. The percentage of correct decisions increased with exposure more rapidly at low dose levels and levelled off at high doses. Depth discrimination is different from a single-object detection task. The relationships among the perceived depth separation, the SNRs of the individual objects and the percentage of correct decisions are still unknown. Further studies are needed to explore how depth discrimination is related to the SNR of the individual objects and whether the depth discrimination task can be predicted by SNR models.

The intraobserver variability was estimated from two repeated readings of the set of test images by one observer. The difference between the two readings for each exposure level and depth separation was calculated. The difference in the percentage of correct decisions between the two readings varied from 0 to 20%, from 0 to 8% and from 0 to 8%, respectively, for the contact, contact-zoom and magnification modes. The variability generally increased with decreasing depth separation and decreasing exposure. There were no differences between the two readings at depth separations of 8 mm and 10 mm for all exposure levels and all image modes. For the magnification mode, the two repeated readings were identical for 18 of the 20 conditions (5 depth separations $\times 4$ exposure levels). A difference was observed only at a 2 mm separation with the two lower exposure levels, 4 mAs and 8 mAs. For both the contact

and contact-zoom modes, 14 of the 20 conditions were identical. Four of the six differences occurred at a 2 mm separation with the four exposure levels (4, 8, 32, 63 mAs).

In a previous study, we designed a virtual cursor to measure the absolute depths of the fibrils in stereo phantom images similar to those used in this study. We found that the accuracy of the absolute depth measurement also depended on the image modes. The accuracy of the measurements in contact-zoom mode was similar to those in contact mode, whereas the accuracy was higher when measured in the images acquired with geometric magnification (Goodsitt *et al* 2002). Although these results are consistent with the observation in the current depth perception experiments, the task of visual depth discrimination and the task of absolute depth measurement are different. The former task requires the differentiation of the relative depths of two objects, whereas, the latter task measures the absolute depth of a single object with a calibrated virtual cursor. We found that an observer who could perform the depth discrimination task very well might fail to make accurate depth measurements. In clinical practice, visual depth discrimination is likely to be the more important task during detection and diagnostic workup. Depth measurement is likely to be more important for treatment planning and intervention.

We used a thin phantom (7 mm) for the experiments in this study. A thin phantom would result in less scattered radiation than a thick phantom (e.g., 4 to 5 cm thick for an average breast). The effect of scattered radiation in digital imaging is mainly a reduction of the SNR for a given exposure. In this study, we reduced the SNR by reducing the exposure. The overall effect of increased scatter on the percentage of correct decision for a given exposure would be similar to that of lower x-ray exposure. Since scattered radiation is reduced by the air gap in the magnification geometry and by the Bucky grid in the contact geometry, the effect of scatter would be relatively small even if a thicker phantom was used. In our previous study (Chan *et al* 2001), we sandwiched the modular phantom between two 2 cm thick BR-12 breast tissue equivalent material to simulate the scatter of an average breast. The relative trend between different exposure conditions observed in the previous study was similar to that observed in the current study. We therefore expect that the relative trend of the different imaging conditions will not be affected by the thickness of the phantom.

In this study, acquisition of the stereo image pair was accomplished by shifting the phantom; there was no grid cutoff due to misalignment of the x-ray beam and the focused grid. In patient imaging, the stereo image pair has to be acquired by shifting the focal spot as shown in figure 1(a). The central ray of the x-ray beam will not be aligned with the central axis of the grid, thus artefacts due to grid cutoff may occur. We have performed a pilot study acquiring stereomammograms of a few patients with Institutional Review Board approval and patient informed consent. We did not observe obvious grid cutoff artefacts in the displayed stereomammograms, probably because the grid ratio is low and the stereo shift is small. If 3D image acquisition methods such as digital tomosynthesis and stereomammography are implemented clinically in the future, we believe that a new grid design should be developed for the digital mammography system so that there will not be a trade-off between SNR and 3D information when using these new imaging techniques.

In this study, we used test objects containing a crossing fibril pair perpendicular to each other with the horizontal fibril parallel to the stereo shift direction. We observed a trend that the depth discrimination was poorer if the horizontal fibril was in front of the crossing pair. In our previous study of absolute depth measurement with a virtual cursor (Goodsitt *et al* 2000), we also found that the accuracy of measuring the depth of horizontal fibrils was much worse than that of vertical fibrils. However, in actual clinical images, there will be few, if any, objects of interest that consist of linear and uniform structures exactly parallel to each other and parallel to the stereo shift direction. Therefore, the angular orientation of the two

objects to be discriminated will not be a major issue in clinical practice as long as they contain structures that are oriented at a moderate angle relative to the stereo shift direction.

5. Conclusions

The accuracy of depth discrimination of fibrils in stereomammography increases with x-ray exposure and depth separation and depends on imaging geometry. At a stereo angle of $\pm 3°$, a 2 mm depth resolution was achieved with over 60% accuracy for all imaging conditions studied. It improved to greater than 90% accuracy at higher doses in contact geometry. With magnification stereomammography, a 2 mm depth resolution was achieved with greater than 90% accuracy for all exposure levels studied. It was found that zooming the contact stereo images by 2X did not improve the accuracy under our experimental conditions. When the images were noisy and the depth separation between the fibrils was small, depth discrimination was significantly better in stereo images acquired with geometric magnification than in images acquired with a contact technique and displayed with or without zooming. These results indicate that stereoscopic imaging, and in particular, magnification stereomammography, may be useful for visualizing the spatial distribution of microcalcifications in a cluster and differentiating overlapping tissues from masses on mammograms. Further studies are underway to investigate the dependence of depth perception on the shape and size of the objects and to evaluate if specially designed cursors can assist in depth discrimination of target objects in stereoscopic images. An observer experiment is being conducted to evaluate the characterization of mammographic lesions in stereo images of biopsy tissue specimens (Chan *et al* 2003).

Acknowledgments

This work is supported by US Army Medical Research and Materiel Command Grants DAMD 17-98-1-8210 and DAMD17-99-1-9294. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred.

References

Aufrichtig R 1999 Comparison of low contrast detectability between a digital amorphous silicon and a screen-film based imaging system for thoracic radiography *Med. Phys.* **26** 1349–58

Beam C A, Layde P M and Sullivan D C 1996 Variability in the interpretation of screening mammograms by US radiologists—findings from a national sample *Arch. Intern. Med.* **156** 209–13

Bird R E, Wallace T W and Yankaskas B C 1992 Analysis of cancers missed at screening mammography *Radiology* **184** 613–7

Boone J M, Nelson T R, Lindfors K K and Seibert J A 2001 Dedicated breast CT: radiation dose and image quality evaluation *Radiology* **221** 657–67

Chan H P, Goodsitt M M, Darner K L, Sullivan J M, Hadjiiski L M, Petrick N and Sahiner B 2001 Effects of stereoscopic imaging technique on depth discrimination *Proc. IWDM 2000—5th Int. Workshop on Digital Mammography, Digital Mammography (Toronto, Canada)* ed M J Yaffe (Madison, WI: Medical Physics Publishing)

Chan H P, Goodsitt M M, Hadjiiski L, Roubidoux M A, Bailey J E, Helvie M A, Lydick J T and Sahiner B 2003 ROC study comparing radiologists' performances in evaluating breast lesions on stereoscopic and single-projection digital specimen mammograms *Med. Phys.* **30** 1456 (abstract)

Conant E F, Maidment A D, Albert M, Piccoli C W, Nussbaum S A and McCue P A 1996 Small field-of-view digital imaging of breast calcifications: method to improve diagnostic specificity *Radiology* **201(P)** 369

Curry T S, Dowdey J E and Murry R C 1992 *Christensen's Physics of Diagnostic Radiology* 4th edn (Philadelphia, PA: Lea and Febiger)

Doi K and Duda E E 1983 Detectability of depth information by use of magnification stereoscopic technique in cerebral angiography *Radiology* **146** 91–5

Doi K and Imhof H 1977 Noise reduction by radiographic magnification *Radiology* **122** 479–87

Doi K, Patronas N J, Duda E E, Geldner E and Dietz K 1981 X-ray imaging of blood vessels to the brain by use of magnification stereoscopic technique *Adv. Neurol.* **30** 175–89

Doi K and Rossmann K 1974 The effect of radiographic magnification on blood vessel imaging with various screen-film systems *Med. Phys.* **1** 257–61

Getty D J, Pickett R M and D'Orsi C J 2001 Stereoscopic digital mammography: improving detection and diagnosis of breast cancer *Computer Assisted Radiology and Surgery 2001, Proc. 15th Int. Congress and Exhibition (Berlin) (International Congress Series 1230)* ed H U Lemke, M W Vannier, K Inamura, A G Farman and K Doi (Amsterdam: Elsevier) pp 506–11

Goodsitt M M, Chan H P, Darner K L and Hadjiiski L M 2002 The effects of stereo shift angle, geometric magnification and display zoom on depth measurements in digital stereomammography *Med. Phys.* **29** 2725–34

Goodsitt M M, Chan H P and Hadjiiski L M 2000 Stereomammography: evaluation of depth perception using a virtual 3D cursor *Med. Phys.* **27** 1305–10

Harvey J A, Fajardo L L and Innis C A 1993 Previous mammograms in patients with impalpable breast carcinomas: retrospective vs blinded interpretation *Am. J. Roentgenol.* **161** 1167–72

Higashida Y, Hirata Y, Saito R, Doudanuki S, Bussaka H and Takahashi M 1988 Depth determination on stereoscopic digital subtraction angiograms *Radiology* **168** 560–2

Jackson V P, Hendrick R E, Feig S A and Kopans D B 1993 Imaging of the radiographically dense breast *Radiology* **188** 297–301

Kelsey C A, Moseley R D, Mettler S A and Briscoe D E 1982 Cost-effectiveness of stereoscopic radiographs in detection of lung nodules *Radiology* **142** 611–3

Loo L N, Doi K and Metz C E 1985 Investigation of basic imaging properties in digital radiography: 4. Effect of unsharp masking on the detectability of simple patterns *Med. Phys.* **12** 209–14

Maidment A D A, Albert M, Conant E F, Piccoli C W and McCue P A 1996 Prototype workstation for 3-D diagnosis of breast calcifications *Radiology* **201(P)** 556

Maidment A D A, Bakic P and Alberg M 2003 Is stereomammography possible without increasing dose? *Proc. IWDM 2002—6th Int. Workshop on Digital Mammography, Digital Mammography (Bremen, Germany)* ed H Petitgen (Berlin: Springer) pp 510–5

Niklason L T *et al* 1997 Digital tomosynthesis in breast imaging *Radiology* **205** 399–406

Ragnarsson J I and Karrholm J 1992 Factors influencing postoperative movement in displaced femoral neck fractures: evaluation by conventional radiography and stereoradiography *J. Orthop. Trauma* **6** 152–8

Raptopoulos V, Baum J K, Hochman M, Karellas A, Houlihan M J and D'Orsi C J 1996 High resolution CT mammography of surgical biopsy specimens *J. Comput. Assist. Tomogr.* **20** 179–84

Seidman H, Gelb S K, Silverberg E, LaVerda N and Lubera J A 1987 Survival experience in the breast cancer detection demonstration project *CA Cancer J. Clin.* **37** 258–90

Simons K 1981 Stereoacuity norms in young children *Arch. Ophthalmol.* **99** 435–45

Suryanarayanan S, Karellas A, Vedantham S, Glick S J, D'Orsi C J, Baker S P and Webber R L 2000 Comparison of tomosynthesis methods used with digital mammography *Acad. Radiol.* **7** 1085–97

Trocme M C, Sather A H and An K N 1990 A biplanar cephalometric stereoradiography technique *Am. J. Orthod. Dentofacial Orthop.* **98** 168–75

Wallis M G, Walsh M T and Lee J R 1991 A review of false negative mammography in a symptomatic population *Clin. Radiol.* **44** 13–5

application specific integral circuit (ASIC) electronics to provide high-count rate and short scanning time. Basic design parameters of the system are 20x20 cm^2 full field, 100x100 μm^2 pixel size, 6 s scanning time, <2% scatter fraction, >80% photon detection efficiency, <2% dead time and <0.1photon/pixel noise equivalent signal. The total width of the slits will be 4 mm. These parameters will provide a signal to noise ratio of 5 for 112 μm microcalcifications using a 200 mR input exposure. The proposed full field digital mammography system can potentially provide improved image quality at reduced patient dose.

[1]Shikhaliev, P. M., *NIM*, **A369**(1), (1996),p.147-50.
[2]Shikhaliev, P. M. and S. Molloi, *NIM*, **A487**, (2002),p.676-85.

TH-C24A-04

Automated Stereo Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists and a Computer

M Goodsitt*, H Chan, J Lydick, C Gandra, M Helvie, J Bailey, M Roubidoux, C Paramagul, B Sahiner, N Petrick, University Michigan, Ann Arbor, MI, and CRDH, Rockville, MD

We are developing an automated stereo spot mammography technique for improved imaging of suspicious dense regions within digital mammograms. A critical element in this technique is the automated detection of regions for spot imaging. Previously, we presented the results of a preliminary observer study comparing regions selected in 800-micron digitized mammograms by different radiologists. We have since modified our display program to show images at higher resolution (200 and 400-micron) and to include a comparison with regions selected by a computer using an in-house developed CAD mass detection program. The new graphical-user-interface was employed by 4 radiologists to select regions in 200 digitized mammograms. Both the radiologist and the computer could select up to 3 regions in each image. We quantified the agreement between the regions selected by the radiologist and the computer by an overlap index, which was the ratio of the intersection of a pair of areas selected by the radiologist and the computer to the area selected by the computer in each mammogram. The selected regions were exhaustively paired and their overlap indices calculated. The overlap indices in each image were then ordered from largest to smallest. The averages of the largest indices in the 200 images for the 4 radiologist-to-computer comparisons ranged from 74% to 95%. The averages for the second largest indices ranged from 18% to 73%. The agreement of the selected spot regions among radiologists and the computer will be discussed in detail. Images illustrating the comparisons and additional statistics will be presented.

TH-C24A-05

ROC Study Comparing Radiologists' Performances in Evaluating Breast Lesions On Stereoscopic and Single-Projection Digital Specimen Mammograms

H-P Chan*, MM Goodsitt, L Hadjiiski, MA Roubidoux, JE Bailey, MA Helvie, JT Lydick, B Sahiner, University of Michigan, Ann Arbor, MI

We are evaluating the effects of stereoscopic imaging on the interpretation of mammographic lesions. For the present investigation, we conducted an observer ROC study to compare radiologists' performance in characterizing breast lesions in biopsy tissue specimens from stereo and single-projection digital mammograms. Stereo and single-projection images of 158 tissue specimens were acquired with a GE 2000D digital mammography system. The images were displayed on a DOME/MegaScan stereo display system and viewed with LCD glasses. A graphical user interface was developed that allowed the user to adjust the image display parameters and to record the BI-RADS assessment and confidence ratings on the presence of microcalcifications and masses, margin clearance, and likelihood of malignancy. A preliminary study was conducted with two MQSA radiologists. For one observer, the viewing of stereo images resulted in improvements in the accuracy of estimating the likelihood of malignancy (from Az= 0.72 to 0.74) and BI-RADS assessment (from Az= 0.68 to 0.73), and increased confidence in judging margin clearance ($p<0.00001$). For the other observer, the effect was the opposite - reduced accuracy in estimating the likelihood of malignancy (from Az= 0.76 to 0.70), reduced BI-RADS assessment (from Az= 0.77 to 0.72) and reduced confidence in judging margin clearance ($p<0.03$). These results appear to indicate that characterization of the lesions is correlated with the observer's confidence in visualizing the details in the stereo images, which may be affected by the observer's stereo acuity. The observer study will continue with additional radiologists, and the results will be presented.

TH-C24A-06

Diagnostic Clinical Benefits of Digital 3-D Mammography

M Lehtimäki*[1], M Pamilo[2], L Raulisto[2], M Roiha[2], M Kalke[1], S Siltanen[1], T Ihamäki[1], (1) Instrumentarium Corporation Imaging Division, Tuusula, Finland, (2) Helsinki University Central Hospital Mammography Department, Helsinki, Finland

The purpose of this study is to find out the impact of 3-dimensional digital mammography imaging following analysis of the abnormal findings of screening mammograms. Over a period of eight months, digital 3-D mammography imaging TACT Tuned Aperture Computed Tomography, digital spot imaging (DSI), screen-film mammography imaging (SFM) and diagnostic film imaging (DFM) examinations were performed on 60 asymptomatic cases. All patients were recalled because it was not possible to exclude the presence of breast cancer on screening films. Abnormal findings on the screening films were non-specific tumor-like parenchymal densities, parenchymal asymmetries or distortions with or without microcalcifications or just microcalcifications. Mammography work-up (film imaging) included spot compression and microfocus magnification views. The 3-D softcopy reading in all cases was done with Delta 32 TACT mammography workstation, while the film images were read using a mammography-specific light box. During the softcopy reading only windowing tools were allowed. The result of this study indicates that the clinical diagnostic image quality of digital 3-D and digital spot images are better than in film images, even in comparison with diagnostic work-up films. Potential advantages are to define if the mammography finding is caused by a real abnormal lesion or by superimposition of normal parenchymal structures, to detect changes in breast tissue which would otherwise be missed, to verify the correct target for biopsies and to reduce the number of biopsies performed.

Supported by Instrumentarium Corporation Imaging Division.

TH-C24A-07

Computer Simulation of X-Ray Mammography Using High Resolution Voxel Phantoms

D Dance*[1], R Hunt[1], P Bakic[2], A Maidment[2], M Sandborg[3], G Alm Carlsson[3], (1) The Royal Marsden Hospital, London, United Kingdom, (2) Hospital of the University of Pennsylvania, Philadelphia, PA, (3) University of Linköping, Linköping, Sweden

A computer model of X-ray mammography has been developed which uses realistic high-resolution voxel phantoms to model the breast. These phantoms have 400-micron voxels and simulate in three dimensions adipose and fibroglandular tissues, Cooper's ligaments, ducts and skin. The Monte Carlo computer model calculates the dose to each tissue and image properties such as energy imparted and noise per pixel and scatter-to-primary ratios (S/P). It allows dosimetry for individual breast tissues, and studies of the dependence of image properties on breast structure. The application of the program is illustrated with calculations using a Mo/Mo spectrum at 28kV for a phantom 50 mm thick. The value of S/P (no grid case) was 0.28 at the chest wall with a maximum of 0.71. Variations in S/P due to tissue inhomogeneity were 25%. Doses (normalised to 100 for adipose tissue) were: fibroglandular tissue 92; ducts 87; skin 284; Cooper's ligaments 147. Calculations of the incident air kerma to average glandular dose conversion factor were made for 10 phantoms. Each had the same thickness, but different spatial distribution of tissues within a fixed central region of average glandularity 70%. The values showed variations of only 1%. It is concluded that the use of a computer model of mammography incorporating a realistic breast has significant advantages over models using homogeneous phantoms for dosimetry and the study of image properties.

Supported by Commission of the European Communities Grant CT2000-0036 and US Department of Defense Grant DAMD 17-98-1-8169.